

Tesis doctoral

**Nuevo modelo de evaluación de la
Interacción Persona-Robot en
entornos industriales:
ITPX-ROBOTS**

Ainhoa Apraiz Iriarte

Directores de Tesis

Ganix Lasa Erle

Maitane Mazmela Etxabe

Tesis dirigida a la obtención del título de DOCTOR
por MONDRAGON UNIBERTSITATEA

Departamento de Mecánica y Producción Industrial
Mondragon Unibertsitatea

Diciembre 2023



Reconocimiento (by): Se permite cualquier explotación de la obra, incluyendo una finalidad comercial, así como la creación de obras derivadas cuya distribución está permitida sin ninguna restricción.

DECLARACIÓN DE ORIGINALIDAD

Yo, Ainhoa Apraiz Iriarte, declaro que esta tesis es fruto de mi trabajo personal, y que no ha sido previamente presentada para obtener otro título o calificación profesional. Las ideas, formulaciones, imágenes, ilustraciones tomadas de fuentes ajenas han sido debidamente citadas y referenciadas.

AGRADECIMIENTOS

Quiero expresar mi agradecimiento al Diseinu Berrikuntza Zentroa (DBZ) de Mondragon Unibertsitatea (MU) y a Mondragon Goi Eskola Politeknikoa (MGEP) por brindarme la oportunidad y los recursos que hicieron posible la realización de esta tesis doctoral. Mi gratitud se extiende a todos los miembros de este equipo, con un reconocimiento especial a mis directores, Ganix y Maitane. Su guía y apoyo han sido fundamentales en mi desarrollo académico y he aprendido muchísimo a su lado. Su disposición para ayudar es inspiradora, extrayendo lo mejor de mí en cada paso.

También quiero agradecer a mis compañeros del departamento de robótica por haber hecho posible la ejecución de los casos de estudio. Asimismo, quiero expresar mi gratitud a todas las personas que participaron en los casos de estudio o en las evaluaciones expertas.

Mi agradecimiento se extiende a Aita y Ama, por todo. A Iñigo por su constante presencia y apoyo diario. A mis hermanos, Adiran y Nerea. A Amuma, la fortaleza en persona. A Izeko Nere e Izeko Ane. A mis amigas.

No puedo dejar de expresar mi gratitud a todas las personas que han contribuido de diversas formas a este proyecto. Cada gesto de apoyo y cada momento compartido han dejado una huella en este viaje académico y personal. A todos y todas, os estoy sinceramente agradecida.

RESUMEN

La industria 5.0 tiene como pilar principal el bienestar de las personas, redefiniendo así el rol de las personas y transformando los actuales puestos de trabajo hacia formatos aún desconocidos. Esta investigación se centra en uno de los desafíos más relevantes para la Industria 5.0: la Interacción Persona-Robot (HRI). Concretamente en aquellos escenarios en los que personas y robots comparten el mismo espacio de trabajo para alcanzar un objetivo común de forma sincronizada, cooperativa o colaborativa.

El número de robots en la industria manufacturera ha aumentado significativamente desde hace varias décadas y, en los últimos años, también han aumentado el número y la variedad de industrias que utilizan robots. Sin embargo, para que los robots se conviertan en aliados de las personas usuarias, deben proporcionar experiencias positivas y adecuadas a su propósito mediante interacciones fluidas y satisfactorias, para que así las personas acepten la nueva tecnología. En este contexto, la Experiencia de Usuario (UX) emerge como un marco fundamental que vincula a las personas con los robots, siendo su evaluación una tarea esencial.

Con el objetivo de asegurar la aceptación exitosa de los robots por parte de las personas, esta investigación presenta el modelo ITPX-ROBOTS (Individual Task Performance Experience – Robots) como una propuesta para evaluar la idoneidad de los sistemas robóticos en relación con las personas y para evaluar el logro de tareas en términos de desempeño, fisiología y percepción. Este modelo permite una evaluación integral de la HRI, fusionando el enfoque hedonómico y pragmático. El modelo proporciona insights valiosos sobre la dinámica de la HRI, contribuyendo a mejorar la eficiencia y la eficacia en el entorno laboral.

Ante este contexto, esta tesis doctoral introduce el marco conceptual Human-Robot Interaction Experience (HRIX), que indica cómo el ajuste entre las características de las entidades involucradas en la interacción (persona, robot y contexto), tiene un impacto en la UX y en sus componentes, que en este entorno de HRI son: la utilidad, la facilidad de uso, el confort, la aprendibilidad y la controlabilidad. Todo ello, tiene unas consecuencias en los factores humanos como la actitud e intención de uso, la confianza, la seguridad percibida y la satisfacción.

Además, el modelo se basa en la taxonomía Human-Robot-Context (HUBOXT), donde se identifican (i) las características de la persona que ejercen influencia en la aceptación de la HRI, (ii) las características del robot que ejercen influencia en la aceptación de la HRI, y (iii) las características del contexto que ejercen influencia en la aceptación de la HRI divididas en las características del micro contexto y las características del macro contexto.

Teniendo en cuenta el marco conceptual y la taxonomía, se propone el modelo ITPX-ROBOTS, que se divide en tres fases, siguiendo un enfoque holístico para las evaluaciones de la UX, en consonancia con el modelo ITPX (Mazmela, 2020). Estas tres etapas abarcan: (i) la evaluación previa, que se enfoca en una evaluación experta mediante heurísticas; (ii) la evaluación durante la interacción, que implica la recopilación de indicadores de desempeño y parámetros fisiológicos durante la ejecución de las tareas con el robot; y, (iii) la evaluación posterior a la interacción, que se centra en la obtención de las percepciones de las personas usuarias acerca de la interacción que han experimentado con el robot.

Para la ejecución de la fase previa, la evaluación heurística emerge como un método de bajo costo y fácil implementación que puede aplicarse de manera iterativa en diversas fases del proceso de diseño. En este contexto, se ha desarrollado HEUROBOX, un nuevo conjunto de heurísticas que agrupa directrices de diseño como una lista lógica de heurísticas para evaluar la interacción entre personas y robots en entornos industriales. HEUROBOX consta de 84

heurísticas para la evaluación básica y 228 heurísticas para la evaluación avanzada, permitiendo adaptar la herramienta a diferentes requisitos industriales. Por último, el conjunto de nuevas heurísticas ha sido validado por expertos mediante el cuestionario System Usability Scale (SUS) y el Analytical Hierarchy Process (AHP).

Para complementar la fase posterior a la interacción, se ha desarrollado el cuestionario HUROY, diseñado para evaluar la UX en interacciones con robots en entornos industriales. Se ha desarrollado utilizando un enfoque empírico para la selección de los ítems con el fin de garantizar la relevancia práctica de los mismos. Mediante un análisis psicométrico donde las personas expertas han evaluado cada uno de los ítems según la importancia, necesidad, relevancia y claridad, se ha validado el cuestionario.

La fase experimental del modelo se ha llevado a cabo mediante dos casos de estudio, uno relacionado con una tarea de ensamblado y el otro con una tarea de desensamblado. Los resultados revelan que el modelo ITPX-ROBOTS ayuda a diagnosticar el éxito de la HRI, proporcionando un procedimiento que facilita el análisis del impacto de los robots en la consecución de tareas y la percepción de usuario. En consecuencia, la tesis presenta un modelo para obtener una comprensión más profunda de la integración de los robots, lo que permitirá optimizar procesos industriales, mejorar la calidad y aumentar la productividad de manera efectiva.

El documento detalla el proceso de desarrollo del modelo ITPX-ROBOTS en varios capítulos. En el primer capítulo, tras contextualizar la temática, se enumeran los objetivos e hipótesis de la investigación, acompañados de la metodología de investigación empleada. El segundo capítulo comprende el marco científico-tecnológico y consta de cinco subcapítulos. Para empezar, se presenta una visión general de la HRI en la literatura, estableciendo los conceptos fundamentales. A continuación, se realiza una revisión sistemática de la literatura sobre modelos de evaluación utilizados para medir la UX y la aceptación tecnológica en la HRI en entornos industriales. Seguido, se actualiza esta revisión sistemática con los nuevos avances identificados en la literatura. Luego, se detallan los métodos y métricas utilizados para evaluar la UX y la aceptación, centrándose en la evaluación psicofisiológica y la percepción del usuario a través de cuestionarios. Finalmente, para acabar con el enmarque científico-tecnológico, se realiza un análisis crítico del estado del arte, discutiendo las posibles contribuciones y oportunidades de investigación identificadas. El tercer capítulo introduce el modelo ITPX-ROBOTS. Aquí, se describen los fundamentos teóricos que sustentan su creación y se definen los elementos clave que componen este nuevo modelo, así como el procedimiento experimental. El cuarto capítulo se dedica al desarrollo y validación de la herramienta de evaluación heurística HEUROBOX. En el quinto capítulo, se presenta el desarrollo y la validación del cuestionario HUROY. En el sexto capítulo, se describen los casos de estudio que validan el modelo en su totalidad. Finalmente, en el último capítulo, se presentan las conclusiones y se corroboran las hipótesis planteadas. Además, se destacan las contribuciones más relevantes de la investigación. Entre las aportaciones significativas de este estudio se incluyen:

- Nuevo modelo para evaluar la Interacción Persona-Robot en entornos industriales denominado ITPX-ROBOTS.
- Nuevo listado de evaluación heurística para la Interacción Persona-Robot en entornos industriales denominado HEUROBOX.
- Nuevo cuestionario para evaluar la percepción de las personas usuarias en la Interacción Persona-Robot en entornos industriales denominado HUROY.

LABURPENA

5.0 Industriaren oinarri nagusia pertsonen ongizatea da. Testuinguru berri honetan, pertsonen rola guztiz aldatzen da, gaur egungo lanpostuak oraindik ezezagunak diren formatuetara eraldatuz. Ikerketa hau 5.0 Industriako erronka garrantzitsuenetako batean oinarritzen da: Pertsonen eta Roboten arteko Interakzioan (HRI), hain zuzen ere. Zehazki, ikerketa honek, pertsonen eta roboten lan-eremu bera partekatzen duten agertokietan du oinarria, bien artean eginkizun komun bat modu sinkronizatuan, kooperatiboan edo lankidetzan lortu beharreko lanetan.

Manufaktura-industrian erabilitako robot kopurua etengabe igo da duela hamarkada batzutatik hona. Gainera, azken urteotan, robotak erabiltzen dituzten industria motek eta lantoki kopuruek ere gorakada nabarmena izan dute. Hala ere, robot hauek langileen egunerokotasunean bidelagun bihurtu daitezke, hauen erabilera-esperientzia positiboa ezinbestekoa da. Horretarako, langileen betebeharrak asetzeko egokituak izan behar dira, elkarreragin arin eta gogobetegarrien bidez. Ezean, teknologia berriak onartzeko zailtasunak emango dira langileen artean. Testuinguru horretan, Erabiltzailearen Esperientzia (UX) alorra, pertsonak robotekin lotzen dituen funtsezko esparru gisa agertzen da. Era berean, UX-aren ebaluazioa funtsezko zeregin bilakatzen da.

Pertsonen robotak arrakastaz onartzen dituztela ziurtatzeko, ikerketa honek ITPX-ROBOTS (Individual Task Performance - Robots) eredu aurkeztu du. Eredu hau, sistema robotikoek pertsonen egokitasuna eta eginkizunen lorpena ebaluatzeko proposamena da. Bertan, jarduna, fisiologia eta pertzepzioa bezalako kontzeptuak balioztatzen dira. Ondorioz, eredu honek HRI-aren ebaluazio integrala ahalbidetzen du, ikuspegi hedonomikoa eta pragmatikoa uztartuz. Ereduak HRI-aren dinamikari buruzko ekarpen baliotsuak ematen ditu, lan-ingurunean eraginkortasuna hobetzen lagunduz.

Horrela izanik, doktorego-tesi honetan Human-Robot Interaction Experience (HRIX) kontzeptu-esparrua garatu da. Honen arabera, elkarrekintzan inplikaturako erakundeen (pertsona, robotak eta testuinguruak) ezaugarrien arteko doikuntzak eragina du UX-ean eta haren osagaietan. Kontzeptu-esparru honetan, UX-aren osagai hauek definitu dira: erabilgarritasuna, erabiltzeko erraztasuna, erosotasuna, ikasteko gaitasuna eta kontrolagarritasuna. Horrek guztiak ondorioak ditu giza faktoreetan, hala nola erabiltzeko jarreran eta asmoan, konfiantzan, hautemandako segurtasunean eta gogobetetzean.

Horrez gain, eredu Human-Robot-Context (HUBOXT) taxonomian oinarritzen da, non hurrengo ezaugarriak identifikatzen diren: (i) HRIa onartzean eragina duten pertsonaren ezaugarriak, (ii) HRIa onartzean eragina duten robotaren ezaugarriak, eta (iii) HRIa onartzean eragina duten testuinguruaren ezaugarriak, mikro testuinguruaren ezaugarrietan eta makro testuinguruaren ezaugarrietan banatuta.

Esparru kontzeptuala eta taxonomia kontuan hartuta, ITPX-ROBOTS eredu garatu da, hiru fasetan banatzen dena. Bertan, UX-aren ebaluazioetarako ikuspegi holistikoa jarraitu da, ITPX ereduarekin bat etorriz (Mazmela, 2020). Hiru etapa horietan honako hauek sartzen dira: (i) alde aurreko ebaluazioa, heuristiken bidezko ebaluazioa aditu baten bitartez gauzatzen dena; (ii) elkarreraginean zehar egindako ebaluazioa, lanak robotarekin egin bitartean jardunaren adierazleak eta parametro fisiologikoak biltzea dakarrena; eta (iii) elkarreraginaren ondorengo ebaluazioa, erabiltzaileek robotarekin izan duten interakzioari buruz dituzten pertzepzioak lortzean zentratzen dena.

Elkarreraginaren aurreko fasea gauzatzeko, ebaluazio heuristikoa metodo aproposa da, kostu txikiko eta inplementazio errazeko metodoa baita. Horrez gain, modu iteratiboan aplikatu daiteke diseinu-prozesuaren hainbat fasetan. Testuinguru horretan, HEUROBOX erraminta

garatu da, industria-inguruneetako pertsonen eta roboten arteko elkarreragina ebaluatzeko heuristika-zerrenda. HEUROBOXek 84 heuristika ditu oinarritzko ebaluaziorako eta 228 heuristika ebaluazio aurreraturako. Honek, tresna hainbat industria-baldintzatarara egokitzeak aukera ematen du. Azkenik, heuristika berrien multzoa adituek baliozkotu dute System Usability Scale (SUS) eta Analytical Hierarchy Process (AHP) galdera-sorten bidez.

Interakzioaren ondorengo fasea osatzeko, HUROY galdetegia garatu da, industria-inguruneetan robotekin egindako elkarrekintzetan UX-a ebaluatzeko diseinatua. Itemak hautatzeko ikuspegi empirikoa erabiliz garatu da, itemen garrantzi praktikoa bermatzeko. Adituek item bakoitza garrantziaren, beharraren, adierazgarritasunaren eta argitasunaren arabera ebaluatu duten analisi psikometriko baten bidez, galdetegia baliozkotu da.

Ereduaren fasea esperimental bi azterketa-kasuren bidez gauzatu da. Lehenengoa, mihizatze-lan batekin lotuta dago. Bigarrena aldiz, desmihizatze-lan batekin du zerikusia. Emaitzek erakusten dutenez, ITPX-ROBOTS ereduak HRI-aren arrakasta diagnostikatzen laguntzen du. Izan ere, robotek zereginen lorpenean eta erabiltzaileen pertzepzioan duten eragina aztertzea errazten duen prozedura bat eskaintzen du. Ondorioz, tesi honek, roboten integrazioa sakonago ulertzeko eredu bat aurkeztu du, eta horrek prozesu industrialak optimizatzea, kalitatea hobetzea eta produktibitatea modu eraginkorrean handitzea ahalbidetuko du.

Dokumentu honek ITPX-ROBOT eredu garatzeko prozesua zehazten du hainbat kapitulutan. Lehenengo kapituluan, gaia testuinguruan kokatu ondoren, ikerketaren helburuak eta hipotesiak zerrendatzen dira, erabilitako ikerketa-metodologiarekin batera. Bigarren kapituluak esparru zientifiko-teknologikoa aztertzen du, eta bost azpikapitulu ditu. Lehenik eta behin, Pertsona-Robot Interakzioaren (HRI) ikuspegi orokorra aurkeztu da literaturan, funtsezko kontzeptuak ezarriz. Gero, UX-a neurtzeko erabilitako ebaluazio-ereduei eta industria-inguruneetako HRI-aren onarpen teknologikoari buruzko literaturaren berrikuspen sistematikoa egiten da. Ondoren, berrikuspen sistematiko hori eguneratu da literaturan identifikatutako aurrerapen berriekin. Jarraian, UX-a eta onarpena ebaluatzeko erabilitako metodoak eta metrikak zehazten dira, ebaluazio psikofisiologikoan eta erabiltzaileen galdetegiaren bidez duen pertzepzioan oinarrituta. Azkenik, bosgarren azpikapituluan, esparru zientifiko-teknologikoaren azterketa kritikoa egiten da, identifikatutako ekarpenak eta ikerketa-aukerak eztabaidatuz. Hirugarren kapituluan ITPX-ROBOT eredu aurkeztu da. Hemen, ereduaren sorkuntzaren oinarri diren oinarri teorikoak deskribatzen dira, eta eredu berri hori osatzen duten funtsezko elementuak definitzen dira, baita prozedura esperimental ere. Laugarren kapituluan, HEUROBOX ebaluazio heuristikorako tresnaren garapena eta baliozkotzea aurkeztu dira. Bosgarren kapituluan, HUROY galdetegiaren garapena eta baliozkotzea aurkeztu da. Seigarren kapituluan, eredu osoa baliozkotzen duten azterketa-kasuak deskribatzen dira. Azken kapituluan, ondorioak aurkeztu eta planteatutako hipotesiak berresten dira. Gainera, ikerketaren ekarpen garrantzitsuenak nabarmentzen dira. Azterlan honen ekarpen esanguratsuen artean honako hauek daude:

- ITPX-ROBOTS izeneko eredu berria, ingurune industrialetan pertsonen eta roboten arteko elkarrekintza ebaluatzeko.
- HEUROBOX izeneko ebaluazio heuristikoa gauzatzeko tresna berria, ingurune industrialetako pertsonen eta roboten arteko interakziorako.
- HUROY izeneko galdetegi berria, ingurune industrialetako pertsonen eta roboten arteko elkarreraginean erabiltzaileek duten pertzepzioa ebaluatzeko.

ABSTRACT

The Industry 5.0 has as its main pillar the well-being of people, thus redefining the role of individuals and transforming current job positions into formats that are still unknown. This research focuses on one of the most relevant challenges for Industry 5.0: Human-Robot Interaction (HRI), specifically in scenarios where people and robots share the same workspace to achieve a common goal in a synchronized, cooperative, or collaborative manner.

The number of robots in the manufacturing industry has significantly increased over the past few decades, and in recent years, the number and variety of industries using robots have also grown. However, for robots to become allies of the users, they must provide positive and purposeful experiences through smooth and satisfactory interactions so that people can accept this new technology. In this context, User Experience (UX) emerges as a fundamental framework that connects people with robots, with its evaluation being an essential task.

With the aim of ensuring the successful acceptance of robots by people, this research introduces the ITPX-ROBOTS model (Individual Task Performance Experience - Robots) as a proposal to assess the suitability of robotic systems in relation to individuals and to evaluate task achievement in terms of performance, physiology, and perception. This model allows for a comprehensive evaluation of Human-Robot Interaction (HRI), merging both hedonic and pragmatic approaches. The model provides valuable insights into the dynamics of HRI, contributing to improving efficiency and effectiveness in the workplace.

In this context, this doctoral thesis introduces the conceptual framework of Human-Robot Interaction Experience (HRIX), which illustrates how the alignment between the characteristics of the entities involved in the interaction (person, robot, and context) impacts the User Experience (UX) and its components. In the HRI environment, these components include utility, ease of use, comfort, learnability, and controllability. All of these factors have consequences for human-related factors such as attitude and intention to use, trust, perceived safety, and satisfaction.

Furthermore, the model is based on the Human-Robot-Context taxonomy (HUBOXT), which identifies (i) the characteristics of the person that influence the acceptance of HRI, (ii) the characteristics of the robot that influence the acceptance of HRI, and (iii) the characteristics of the context that influence the acceptance of HRI, divided into micro-context and macro-context characteristics.

Taking into account the conceptual framework and the taxonomy, the ITPX-ROBOTS model is proposed, which is divided into three phases, following a holistic approach to UX evaluations, in line with the ITPX model (Mazmela, 2020). These three stages encompass: (i) the pre-interaction evaluation, which focuses on expert evaluation using heuristics; (ii) the evaluation during the interaction, which involves collecting performance indicators and physiological parameters during the execution of tasks with the robot; and (iii) the post-interaction evaluation, which focuses on obtaining user perceptions about the interaction they have experienced with the robot.

To execute the pre-interaction phase, heuristic evaluation emerges as a low-cost and easily implementable method that can be applied iteratively in various stages of the design process. In this context, HEUROBOX has been developed, a new set of heuristics that groups design guidelines into a logical list of heuristics for evaluating the interaction between people and robots in industrial environments. HEUROBOX consists of 84 heuristics for basic evaluation and 228 heuristics for advanced evaluation, allowing the tool to be adapted to different industrial requirements. Finally, the set of new heuristics has been validated by experts using the System Usability Scale (SUS) questionnaire and the Analytical Hierarchy Process (AHP).

To complement the post-interaction phase, the HUROY questionnaire has been developed, designed to evaluate the User Experience (UX) in interactions with robots in industrial environments. It has been developed using an empirical approach for item selection to ensure their practical relevance. Through psychometric analysis, where experts have evaluated each of the items for importance, necessity, relevance, and clarity, the questionnaire has been validated.

The experimental phase of the model has been carried out through two case studies, one related to an assembly task and the other to a disassembly task. The results reveal that the ITPX-ROBOTS model helps diagnose the success of HRI, providing a procedure that facilitates the analysis of the impact of robots on task accomplishment and user perception. Consequently, the thesis presents a model to gain a deeper understanding of the integration of robots, which will enable the optimization of industrial processes, improve quality, and increase productivity effectively.

The document details the development process of the ITPX-ROBOTS model in several chapters. In the first chapter, after providing context, the research objectives and hypotheses are listed, along with the research methodology employed. The second chapter encompasses the scientific and technological framework and consists of five sub-chapters. Firstly, it provides an overview of HRI in the literature, establishing fundamental concepts. Next, it conducts a systematic literature review of evaluation models used to measure UX and technological acceptance in HRI in industrial environments. Following that, this systematic review is updated with new advancements identified in the literature. Then, it details the methods and metrics used to assess UX and acceptance, focusing on psychophysiological evaluation and user perception through questionnaires. Finally, to conclude the scientific and technological framework, a critical analysis of the state of the art is conducted, discussing possible contributions and identified research opportunities. The third chapter introduces the ITPX-ROBOTS model. Here, it describes the theoretical foundations that underpin its creation and defines the key elements that make up this new model, as well as the experimental procedure. The fourth chapter is dedicated to the development and validation of the heuristic evaluation tool HEUROBOX. In the fifth chapter, the development and validation of the HUROY questionnaire are presented. In the sixth chapter, the case studies validating the model in its entirety are described. Finally, in the last chapter, conclusions are presented, and the hypotheses are confirmed. Furthermore, the most relevant contributions of the research are highlighted. Some of the significant contributions of this study include:

- A new model for evaluating Human-Robot Interaction in industrial environments, named ITPX-ROBOTS.
- A new list of heuristic evaluation criteria for Human-Robot Interaction in industrial environments, named HEUROBOX.
- A new questionnaire for assessing user perceptions in Human-Robot Interaction in industrial environments, named HUROY.

PUBLICACIONES VINCULADAS A ESTA TESIS DOCTORAL

Redacción y envío de artículos para su revisión y contribución

2023:

- **Apraiz, A.**, Lasa, G. & Mazmela, M. Evaluation of User Experience in Human-Robot Interaction: A Systematic Literature Review. International Journal of Social Robotics (2023). <https://doi.org/10.1007/s12369-022-00957-z>
- **Apraiz, A.**, Mulet Alberola, J.A., Lasa, G., Mazmela, M., & Ngoc Nguyen, H. Development of a new set of Heuristics for the evaluation of Human-Robot Interaction in industrial settings: Heuristics Robots Experience (HEUROBOX). Frontiers in Robotics (2023). AI 10:1227082. <https://doi.org/10.3389/frobt.2023.1227082>
- **Apraiz, A.**, Lasa, G., Montagna, F., Blandino, G., Triviño-Tonato, E. & Dacal-Nieto, A. An Experimental Protocol for Human Stress Investigation in Manufacturing Contexts: Its Application in the NO-STRESS Project. Systems (2023), 11, 448. <https://doi.org/10.3390/systems11090448>

2021:

- **Apraiz-Iriarte, A.**, Lasa-Erle, G., Mazmela-Etxabe, M. Application of questionnaires for evaluation of user experience in industrial environments. DYNA (2021), 96(3). 235. <https://doi.org/10.6036/10025>
- **Apraiz-Iriarte, A.**, Lasa-Erle, G., Mazmela-Etxabe, M. Present and future of physiological monitoring to evaluate user experience. DYNA (2021), 96(4). 342. <https://doi.org/10.6036/10218>
- Mazmela-Etxabe, M., Lasa-Erle, G., **Apraiz-Iriarte, A.** Interface design and its impact on software appropriation. DYNA (2021), 96(1), 16. <http://dx.doi.org/10.6036/9939>
- **Apraiz-Iriarte, A.**, Lasa-Erle, G., Mazmela-Etxabe, M. Evaluating user experience with physiological monitoring: a systematic literature review. DYNA New Technologies (2021), 8(1). [20 p.]. <https://doi.org/10.6036/NT10072>

2020:

- **Apraiz-Iriarte, A.**, Lasa-Erle, G. Questionnaires for evaluation of user experience and usability in industrial environments: bibliographic review. DYNA New Technologies (2020), 7(1). [21 p.]. <https://doi.org/10.6036/NT9845>
- Mazmela-Etxabe, M., Lasa-Erle, G., **Apraiz-Iriarte, A.** Multimethod approach to measure the design factor impact on the appropriation of a competitive intelligence interface. DYNA New Technologies (2020), 7(1). [13 p.]. <https://doi.org/10.6036/NT9805>

Divulgación complementaria a través de congresos:

2023:

- **Apraiz, A.,** Lasa, G., Mazmela, M., Escallada, O., & González de Heredia, A. (2023). Development and validation of Human-Robot eXperience (HUROX) questionnaire for industrial collaborative contexts. 27th International Congress on Project Management and Engineering. Donostia (Spain).

2022:

- **Apraiz Iriarte, A.,** Lasa Erle, G., Serrano Muñoz, A., Elguea Aguinaco, Í., & Arana Arexolaleiba, N. (2022). Evaluation of the user experience of an industrial robotic environment in virtual reality. 26th International Congress on Project Management and Engineering. Terrassa (Spain).
- **Apraiz Iriarte, A.,** Lasa, G., & Mazmela Etxabe, M. (2022). Evaluación heurística de la interacción persona-robot en entornos industriales. Congreso Internacional de Interacción Persona-Ordenador (Interacción). Teruel (Spain).

2021:

- **Apraiz Iriarte, A.,** Aztiria Mendizabal, A., & Lasa Erle, G. (2021). Robot-aren itxura estetikoak eta erabiltzaileen preferentziak. Ikergazte. Vitoria-Gasteiz (Spain). DOI: <https://dx.doi.org/10.26876/ikergazte.iv.03.05>
- **Apraiz Iriarte, A.,** Lasa Erle, G., & Mazmela Etxabe, M. (2021) User preferences and associations with light or dark interfaces. 25th International Congress on Project Management and Engineering. Alcoy (Spain).

ÍNDICE

1	INTRODUCCIÓN	2
1.1	OBJETO DE LA TESIS	2
1.2	CONTEXTO DE LA TESIS	2
1.2.1	<i>La industria hoy</i>	<i>2</i>
1.2.2	<i>Interacción persona-robot.....</i>	<i>5</i>
1.2.3	<i>La importancia de la Aceptación Tecnológica y la Experiencia de Usuario en la HRI</i>	<i>7</i>
1.2.4	<i>Antecedentes en el DBZ.....</i>	<i>13</i>
1.2.5	<i>Interés del tema.....</i>	<i>16</i>
1.3	HIPÓTESIS DE LA TESIS	17
1.4	OBJETIVOS DE LA TESIS	18
1.5	METODOLOGÍA DE LA INVESTIGACIÓN	19
1.6	ESTRUCTURA DE LA TESIS.....	21
2	ENMARQUE CIENTÍFICO-TECNOLÓGICO.....	23
2.1	HRI EN LA NUEVA INDUSTRIA	23
2.1.1	<i>Relaciones persona-robot.....</i>	<i>24</i>
2.2	MODELOS DE EVALUACIÓN EN HRI: REVISIÓN SISTEMÁTICA DE LA LITERATURA (SLR)	30
2.2.1	<i>Metodología.....</i>	<i>32</i>
2.2.2	<i>Resultados.....</i>	<i>37</i>
2.2.3	<i>Resultados a la pregunta de investigación 1: ¿Existe un modelo de evaluación del HRI que incluya los factores humanos en entornos industriales?</i>	<i>40</i>
2.2.4	<i>Resultados a la pregunta de investigación 2: ¿Qué factores humanos incluye y cómo los evalúa?.....</i>	<i>73</i>
2.3	ÚLTIMOS AVANCES EN LA LITERATURA	79
2.3.1	<i>Marcos teóricos</i>	<i>79</i>
2.3.2	<i>Herramientas.....</i>	<i>89</i>
2.3.3	<i>Estudios experimentales.....</i>	<i>99</i>
2.4	MÉTODOS Y MÉTRICAS PARA EVALUAR LA UX Y LA ACEPTACIÓN TECNOLÓGICA.....	103
2.4.1	<i>Evaluación psicofisiológica</i>	<i>104</i>
2.4.2	<i>Evaluación de la percepción de usuario</i>	<i>109</i>
2.5	ESTUDIO CRÍTICO DEL ESTADO DEL ARTE	113
2.5.1	<i>Factores humanos influyentes en la aceptación de los robots por parte de las personas en los contextos industriales.....</i>	<i>113</i>
2.5.2	<i>Métodos y medidas de evaluación de los factores humanos que influyen en la aceptación de la HRI en entornos industriales.....</i>	<i>120</i>
2.5.3	<i>La influencia de las variables cinemáticas en la percepción de las personas usuarias</i>	<i>129</i>
2.5.4	<i>Optimización de la UX en la HRI a través del feedback de las interfaces</i>	<i>130</i>
3	MODELO ITPX-ROBOTS.....	136
3.1	FUNDAMENTOS TEÓRICOS	136
3.1.1	<i>Marco conceptual HRIX.....</i>	<i>137</i>
3.1.2	<i>Factores influyentes y métricas en la HRI.....</i>	<i>139</i>
3.1.3	<i>De los factores influyentes a la evaluación.....</i>	<i>148</i>
3.2	MODELO DE EVALUACIÓN ITPX-ROBOTS	149
3.2.1	<i>Evaluación experta.....</i>	<i>151</i>
3.2.2	<i>Evaluación DURANTE la ejecución de tareas.....</i>	<i>152</i>
3.2.3	<i>Evaluación POSTERIOR a la ejecución de las tareas.....</i>	<i>154</i>
4	DESARROLLO Y VALIDACIÓN DE LA HERRAMIENTA DE EVALUACIÓN HEURÍSTICA HEUROBOX..	158
4.1	DIMENSIONES A CONSIDERAR PARA UN NUEVO LISTADO DE HEURÍSTICOS	159
4.2	METODOLOGÍA	160
4.2.1	<i>Fase 1: Búsqueda de la literatura</i>	<i>160</i>

4.2.2	<i>Fase 2: Selección de artículos</i>	161
4.2.3	<i>Fase 3: Selección, agrupación y primera definición heurística</i>	162
4.2.4	<i>Fase 4: Balance y consistencia</i>	164
4.2.5	<i>Fase 5: Validación y priorización</i>	164
4.3	LISTADO COMPLETO DE HEURÍSTICOS	165
4.3.1	<i>Seguridad</i>	166
4.3.2	<i>Ergonomía</i>	167
4.3.3	<i>Funcionalidad</i>	167
4.3.4	<i>Interfaces</i>	168
4.4	RESULTADOS DE LA VALIDACIÓN Y PRIORIZACIÓN DE LAS PERSONAS EXPERTAS	169
4.4.1	<i>Evaluación de la usabilidad</i>	170
4.4.2	<i>Evaluación AHP</i>	171
4.5	EL INDICADOR DE HEUROBOX	173
4.5.1	<i>Identificación de heurísticos cumplidos</i>	174
4.5.2	<i>Cálculo de los puntajes parciales por categoría</i>	174
4.5.3	<i>Cálculo del indicador HEUROBOX</i>	175
4.6	DISCUSIÓN Y CONCLUSIONES	176
4.6.1	<i>La adecuación de HEUROBOX a los requisitos</i>	176
4.6.2	<i>Evaluación sobre HEUROBOX</i>	177
4.6.3	<i>Fortalezas y limitaciones de HEUROBOX</i>	178
5	DESARROLLO Y VALIDACIÓN DEL CUESTIONARIO HUMAN-ROBOT EXPERIENCE (HUROX)	182
5.1	METODOLOGÍA	182
5.1.1	<i>Fase 1: Definición del marco conceptual</i>	183
5.1.2	<i>Fase 2: Definición de los ítems iniciales</i>	183
5.1.3	<i>Fase 3: Medidas de validación</i>	185
5.1.4	<i>Fase 4: Ratio de validez del contenido (CVR)</i>	186
5.1.5	<i>Fase 5: Índice de validez del contenido (CVI)</i>	186
5.1.6	<i>Caracterización de las personas expertas</i>	187
5.2	RESULTADOS	187
5.2.1	<i>Validez aparente del cuestionario</i>	187
5.2.2	<i>Ratio de Validez de Contenido del Cuestionario</i>	188
5.2.3	<i>Índice de Validez de Contenido del Cuestionario</i>	188
5.3	REFINAMIENTO DE LOS ÍTEMS PARA EVALUAR LA HRI EN CONTEXTOS INDUSTRIALES	194
5.3.1	<i>Utilidad percibida</i>	194
5.3.2	<i>Facilidad de uso percibida</i>	194
5.3.3	<i>Seguridad percibida</i>	195
5.3.4	<i>Controlabilidad</i>	195
5.3.5	<i>Aprendibilidad</i>	196
5.3.6	<i>Actitud</i>	196
5.3.7	<i>Satisfacción</i>	196
5.4	DISCUSIÓN Y CONCLUSIONES	197
6	VALIDACIÓN DEL MODELO ITPX-ROBOTS	200
6.1	ENTORNO Y EQUIPAMIENTO UTILIZADO	201
6.1.1	<i>Entorno de los casos de estudio</i>	201
6.1.2	<i>Robot utilizado: KUKA LBR iiwa 14 R820</i>	201
6.1.3	<i>Dispositivos fisiológicos</i>	201
6.2	CASO DE ESTUDIO 1: INFLUENCIA DE LAS VARIABLES CINEMÁTICAS	202
6.2.1	<i>Motivación</i>	203
6.2.2	<i>Metodología</i>	203
6.2.3	<i>Resultados</i>	209
6.2.4	<i>Discusión y conclusiones</i>	227
6.3	CASO DE ESTUDIO 2: IMPACTO DE LAS INTERFACES UNIMODALES VS. MULTIMODALES	234
6.3.1	<i>Motivación</i>	234
6.3.2	<i>Metodología</i>	235
6.3.3	<i>Resultados</i>	242

6.3.4	<i>Discusión</i>	267
7	CONCLUSIONES Y LÍNEAS FUTURAS	275
7.1	CONCLUSIONES GENERALES	275
7.2	VALIDACIÓN DE LAS HIPÓTESIS	284
7.3	APORTACIONES MÁS RELEVANTES	286
7.4	LIMITACIONES DE LA INVESTIGACIÓN	287
7.5	LÍNEAS FUTURAS DE LA INVESTIGACIÓN	289
7.5.1	<i>Optimizar la herramienta HEUROBOX</i>	290
7.5.2	<i>Validar la fiabilidad de HUROY</i>	291
7.5.3	<i>Identificación de las variables cinemáticas óptimas para la HRI en entornos industriales y la robótica afectiva</i>	291
7.5.4	<i>Medición de la UX Acumulativa y estudios longitudinales</i>	292
7.5.5	<i>Análisis de factores contextuales en casos de estudio más diversificados</i>	293
7.5.6	<i>La ética y la accesibilidad en la HRI</i>	293
8	BIBLIOGRAFÍA	296
	ANEXO A: CUESTIONARIO HUROY	320
	ANEXO B: LISTADO DE HEURÍSTICOS HEUROBOX	328

ÍNDICE DE FIGURAS

Figura 1: El modelo HRCAM. Adaptado de Bröhl et al. (2019).	8
Figura 2: Factores de calidad más representativos analizados en la literatura sobre HRI. Adaptado de Coronado et al. (2022).....	10
Figura 3: Pirámide hedonómica de Hancock. Adaptado de Coronado et al. (2022).	11
Figura 4: Los intervalos temporales de la experiencia de la persona usuaria, los términos para describir el tipo de UX relacionada con los intervalos, y el proceso interno que tiene lugar en los distintos periodos. Adaptado de Roto et al. (2011).	12
Figura 5: Metodología UCAD. Adaptado de Aranburu Zabalo et al. (2017).	13
Figura 6: Metodología DEMOX. Adaptado de Aranburu Zabalo (2020).	14
Figura 7: Herramienta XC. Adaptado de Aranburu et al. (2020).	14
Figura 8: Modelo ITPX. Adaptado de Mazmela Etxabe (2020)	15
Figura 9: Metodología de investigación planteada siguiendo el modelo DIR (Horváth, 2007).	20
Figura 10: Esquema de clasificación de la colaboración entre personas y robots en función de la multiplicidad de los agentes. Adaptado de X.V. Wang et al. (2017).	26
Figura 11: Esquema de clasificación de la colaboración entre personas y robots en función de la iniciativa. Adaptado de X.V. Wang et al. (2017).	27
Figura 12: Interacción y nivel de riesgo de los distintos tipos de relación entre personas y robots. Adaptado de X.V. Wang et al. (2018).	29
Figura 13: Marco para clasificar la HRI: niveles de interacción y relaciones con los distintos tipos de interfaz. Adaptado de Prati et al. (2021).	30
Figura 14: Proceso seguido en la Revisión Sistemática de la Literatura (SLR).	37
Figura 15: El modelo modificado de las siete etapas de la acción. Adaptado de Lindblom & Alenljung (2020).	42
Figura 16: Factores identificados que influyen en la aceptación de la HRC. Adaptado de Meissner et al. (2020).	43
Figura 17: Visualización general de los resultados de Villani et al. (2018).	45
Figura 18: El marco de WOI 4.0 y su propuesta de arquitectura. Adaptado de Cohen et al. (2018).	48
Figura 19: El flujo de trabajo de diseño UX propuesto para la HRI. Adaptado de Prati et al. (2021).	51
Figura 20: Esquema para los profesionales que indica los factores humanos organizativos clave en diferentes TRL. Adaptado de Charalambous et al. (2017).	52
Figura 21: Identificación de categorías, clústeres y subclústeres. Adaptado de Gualtieri et al. (2021).	81
Figura 22: Índice de Calidad para medir la HRC. Adaptado de Kokotonis et al. (2023).	95
Figura 23: Tamaño de la muestra según la herramienta fisiológica.	108
Figura 24: Marco conceptual HRIX.	138
Figura 25: Visión general de la taxonomía HUBOXT.	139
Figura 26: Modelo conceptual de la taxonomía HUBOXT para elaborar el modelo de evaluación.	140
Figura 27: Características de la persona que influyen en la aceptación de la HRI.	141
Figura 28: Características del robot que influyen en la aceptación de la HRI.	143
Figura 29: Características del contexto que influyen en la aceptación de la HRI.	146

Figura 30: HRI FIT - El ajuste entre las características de la persona, las características del robot y las características del contexto.....	148
Figura 31: Marco teórico de la evaluación.....	148
Figura 32: Modelo ITPX-Robots.....	150
Figura 33: Constructos que definen la evaluación experta.....	152
Figura 34: Evaluación DURANTE la ejecución de las tareas en el modelo ITPX-ROBOTS.....	153
Figura 35: Evaluación POSTERIOR a la ejecución de las tareas en el modelo ITPX-ROBOTS.....	155
Figura 36: Metodología para el desarrollo de un nuevo método de evaluación heurística de la HRI.	160
Figura 37: Las puntuaciones obtenidas utilizando el cuestionario SUS para HEUROBOX.....	170
Figura 38: Esquema visual de las ponderaciones para calcular el Indicador HEUROBOX.	174
Figura 39: Metodología seguida para el desarrollo del nuevo cuestionario HUOX.....	183
Figura 40: Robot utilizado en los casos de estudio: KUKA IIWA 14 R820 (<i>LBR iiwa / KUKA AG, s. f.</i>).....	201
Figura 41: Diadem de Bitbrain (2019).....	202
Figura 42: Ring de Bitbrain (2019).....	202
Figura 43: Procedimiento experimental seguido en el caso de estudio 1.....	204
Figura 44: Procedimiento experimental ANTES de la ejecución de tareas en el caso de estudio 1.	204
Figura 45: Procedimiento experimental DURANTE la ejecución de las tareas en el caso de estudio 1. ...	205
Figura 46: Procedimiento experimental DESPUÉS de la ejecución de la tarea en el caso de estudio 1...	206
Figura 47: Tarea del robot y contexto.....	206
Figura 48: Secuencia de tareas en un ensamblaje del caso de estudio 1.	207
Figura 49: Resultados de los indicadores fisiológicos en el procedimiento LR en el caso de estudio 1...	213
Figura 50: Resultados de los indicadores fisiológicos en el procedimiento LR en el caso de estudio 1 desagregados por género.....	215
Figura 51: Resultados de los indicadores fisiológicos en el procedimiento RL en el caso de estudio 1...	216
Figura 52: Resultados de los indicadores fisiológicos en el procedimiento RL en el caso de estudio 1 desagregados por género.....	218
Figura 53: Resultados comparativos del cuestionario para el procedimiento LR.....	220
Figura 54: Resultados comparativos del cuestionario en el procedimiento LR desagregados por género.	221
Figura 55: Resultados comparativos del cuestionario para el procedimiento RL en el caso de estudio 1.	222
Figura 56: Resultados comparativos del cuestionario en el procedimiento RL desagregados por género en el caso de estudio 1.....	223
Figura 57: Representación gráfica de la "campana de velocidad en HRI".....	231
Figura 58: Representación gráfica de la paradoja de la aprendibilidad en la HRI.	233
Figura 59: Procedimiento experimental en el caso de estudio 2.	235
Figura 60: Secuencia de las tareas en el caso de estudio 2.....	237
Figura 61: <i>Set-up</i> de la tarea en el caso de estudio 2.	238
Figura 62: Paso a paso del desensamblado del componente electrónico en el caso de estudio 2.	238
Figura 63: Ejemplos de la interfaz utilizada en el caso de estudio 2.....	239

Figura 64: Personas participantes realizando el test en el caso de estudio 2.	242
Figura 65: Resultados de los indicadores fisiológicos en el procedimiento I en el caso de estudio 2.	245
Figura 66: Resultados de los indicadores fisiológicos desagregados por género en el procedimiento I en el caso de estudio 2.....	246
Figura 67: Resultados de los indicadores fisiológicos en el procedimiento II en el caso de estudio 2.....	247
Figura 68: Resultados de los indicadores fisiológicos desagregados por género en el procedimiento II en el caso de estudio 2.....	248
Figura 69: Resultados comparativos del cuestionario HUROY para el procedimiento I.	250
Figura 70: Resultados comparativos del cuestionario HUROY para el procedimiento I desagregados por género.	252
Figura 71: Resultados comparativos del cuestionario HUROY para el procedimiento II.	253
Figura 72: Resultados comparativos del cuestionario HUROY para el procedimiento II desagregados por género.	255
Figura 73: Resultados comparativos del cuestionario HUROY para la interfaz unimodal desagregados por procedimientos.	257
Figura 74: Resultados comparativos del cuestionario HUROY para la interfaz multimodal desagregados por procedimientos.....	259
Figura 75: Aplicación posible en un estudio longitudinal.....	292

ÍNDICE DE TABLAS

Tabla 1: Diferencias entre la industria 4.0 y la industria 5 (Coronado et al., 2022).....	3
Tabla 2: Clasificación de la HRI en términos de espacio de trabajo, tiempo de trabajo, objetivo y contacto. Adaptado de Schmidtler et al. (2015).....	26
Tabla 3: Clasificación según las características de las relaciones persona-robot.	28
Tabla 4: Preguntas de investigación para la SLR.....	32
Tabla 5: Bases de datos utilizadas en la SLR.....	33
Tabla 6: Ecuación de búsqueda utilizada en cada una de las bases de datos.	34
Tabla 7: Criterios de inclusión y exclusión utilizados para la selección de la literatura.	35
Tabla 8: Preguntas para la evaluación de la calidad.	36
Tabla 9: Puntuaciones de los artículos seleccionados.	36
Tabla 10: Número de artículos identificados por año.....	38
Tabla 11: Revistas en las que más artículos se han identificado.	38
Tabla 12: Impacto de las revistas.	39
Tabla 13: Artículos más citados.....	39
Tabla 14: Clasificación de los artículos identificados por tipología.	40
Tabla 15: Descripción general del procedimiento de reconocimiento de acciones e intenciones en la HRI, fase por fase. Adaptado de Lindblom & Alenljung (2020).	41
Tabla 16: Ítems del cuestionario SE-HRI. Adaptado de von der Pütten & Bock (2018).	53
Tabla 17: Clasificación general de los requisitos de las directrices propuestas. Adaptado de Qbilat et al. (2021).....	54
Tabla 18: Tamaño de la muestra de los estudios experimentales.....	69
Tabla 19: Tipo de robot utilizado en cada estudio experimental.	70
Tabla 20: Visión general y comparativa de los estudios experimentales identificados mediante la SLR...	72
Tabla 21: Factores humanos evaluados y las herramientas utilizadas.	77
Tabla 22: Clasificación de las herramientas identificadas.....	78
Tabla 23: Categorización orientada al rendimiento para las métricas obtenidas por Coronado et al. (2022).	83
Tabla 24: Factores humanos en la era de los cobots. Adaptado de (Faccio et al., 2022).....	88
Tabla 25: Flujo de trabajo de la estrategia de mapeo UX. Fuente: Khamaisi, Brunzini, et al. (2022).	94
Tabla 26: Evaluación para medir el HRC-QI. Adaptado de: Kokotinis et al. (2023).....	97
Tabla 27: Análisis comparativo de los estudios que adoptan un enfoque experimental para analizar los factores humanos en la HRI industrial.....	122
Tabla 28: Bases de datos seleccionadas para la búsqueda de listados de heurísticos o guías en la literatura.....	161
Tabla 29: Ecuación de búsqueda por cada una de las bases de datos.....	161
Tabla 30: Identificación inicial de los heurísticos y guías.	162
Tabla 31: Definición de categorías para la clasificación de directrices de diseño y heurísticas.	163
Tabla 32: Literatura heurística específica existente para equilibrar la nueva lista.....	164

Tabla 33: Presentación del método de evaluación AHP.	165
Tabla 34: El conjunto de heurísticas que componen HEUROBOX.	166
Tabla 35: Listado de expertos que participó en la validación de HEUROBOX.	169
Tabla 36: Pesos de importancia agregados en cada grupo de HEUROBOX de acuerdo con AHP.	171
Tabla 37: La adecuación de HEUROBOX a los requisitos establecidos.	176
Tabla 38: Fuente de los ítems para el cuestionario HUROY.	184
Tabla 39: Constructos iniciales y su definición del cuestionario inicial.	184
Tabla 40: Caracterización de las personas expertas que participaron en la validación del cuestionario HUROY.	187
Tabla 41: Resultados de los cálculos de IIS, CVR y CVI de los ítems iniciales del cuestionario.	189
Tabla 42: Modificaciones realizadas en el constructo de Utilidad Percibida del cuestionario HUROY.	194
Tabla 43: Modificaciones realizadas en el constructo de Facilidad de Uso Percibida del cuestionario HUROY.	194
Tabla 44: Modificaciones realizadas en el constructo de Seguridad percibida y Ergonomía del cuestionario HUROY.	195
Tabla 45: Modificaciones realizadas en el constructo de Controlabilidad del cuestionario HUROY.	195
Tabla 46: Modificaciones realizadas en el constructo de Aprendibilidad del cuestionario HUROY.	196
Tabla 47: Modificaciones realizadas en el constructo de Actitud del cuestionario HUROY.	196
Tabla 48: Modificaciones realizadas en el constructo de Satisfacción del cuestionario HUROY.	197
Tabla 49: Descripción de los casos de estudio.	200
Tabla 50: Resumen de las variables del caso de estudio 1.	209
Tabla 51: Resultado del indicador HEUROBOX en el caso de estudio 1.	210
Tabla 52: Indicadores de desempeño en el procedimiento LR.	211
Tabla 53: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR.	211
Tabla 54: Indicadores de desempeño en el procedimiento RL.	212
Tabla 55: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento RL.	212
Tabla 56: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR.	214
Tabla 57: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR desagregado por género.	215
Tabla 58: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento RL.	216
Tabla 59: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento RL desagregado por género.	218
Tabla 60: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR.	220
Tabla 61: Resultados de la prueba T-test para determinar si existen diferencias significativas en las tareas desagregando por género en el Procedimiento LR.	221
Tabla 62: Resultados de la prueba T-test para determinar si existen diferencias significativas en el procedimiento RL.	222

Tabla 63: Resultados de la prueba T-test para determinar si existen diferencias significativas en las tareas desagregando por género en el Procedimiento RL.	224
Tabla 64: Resultados T-test por ítem (diferencias significativas por género, procedimiento y edad) considerando $\alpha = 0,05$	226
Tabla 65: Resultados generales y consecuencias en base a los indicadores fisiológicos.	228
Tabla 66: Resultados generales y consecuencias en base a los indicadores perceptuales.	228
Tabla 67: Resultados T-test por indicador recogido (diferencias significativas por procedimiento, género y edad) considerando $\alpha = 0,05$	230
Tabla 68: Niveles de experiencia previa con robot colaborativos. Fuente: Gervasi et al. (2022).	236
Tabla 69: Resumen de las variables del caso de estudio 2.	240
Tabla 70: Descripción de la muestra que ha participado en el Caso de Estudio 2.	241
Tabla 71: Resultado del indicador HEUROBOX en el caso de estudio 2.	242
Tabla 72: Ratios de producción en el Procedimiento I.	243
Tabla 73: Ratios de producción en el Procedimiento II.	244
Tabla 74: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento I.	245
Tabla 75: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento I desagregados por género.	246
Tabla 76: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento II.	247
Tabla 77: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento II desagregados por género.	248
Tabla 78: Resultados de la prueba T-test para determinar si existen diferencias significativas en el procedimiento I.	250
Tabla 79: Resultados de la prueba T-test para determinar si existen diferencias significativas entre géneros en el Procedimiento I.	252
Tabla 80: Resultados de la prueba T-test para determinar si existen diferencias significativas entre géneros en el Procedimiento II.	254
Tabla 81: Resultados de la prueba T-test para determinar si existen diferencias significativas entre géneros en el Procedimiento II.	255
Tabla 82: Resultados de la prueba T-test para determinar si existen diferencias significativas entre procedimientos al valorar la interfaz Unimodal.	257
Tabla 83: Resultados de la prueba T-test para determinar si existen diferencias significativas entre procedimientos al valorar la interfaz multimodal.	259
Tabla 84: Resultados T-test por ítem (diferencias significativas por género, procedimiento, experiencia y edad) considerando $\alpha = 0,05$. Se resaltan en negrita las diferencias significativas.	264
Tabla 85: Resultados T-test por indicador recogido (diferencias significativas por procedimiento, género, edad y experiencia) considerando $\alpha = 0,05$. Se resaltan en negrita las diferencias significativas. ...	269
Tabla 86: Correlación entre el indicador HEUROBOX y los indicadores fisiológicos.	284

ABREVIATURAS

Abreviatura	Significado
AHP	Proceso Analítico Jerárquico (Analytic Hierarchy Process)
BVP	Pulso de volumen sanguíneo (Blood Volume Pulse)
CAD	Diseño Asistido por Ordenador (Computer Aided Design)
CH	Calidad Hedónica
CP	Calidad Pragmática
CVI	Índice de Validez del Contenido (Content Validity Index)
CVR	Ratio de Validez del Contenido (Content Validity Ratio)
DBZ	Centro de Innovación en Diseño (Diseinu Berrikuntza Zentroa)
DEMOX	Diseño y Evaluación de la Experiencia Máquina-Operador (Design and Evaluation for Machine-Operator eXperiences)
DIR	Diseño de Investigación Inclusiva (Design Inclusive Research)
DSA	Desarrollo de Software Ágil
EC	Evaluación de Calidad
ECG	Electrocardiograma
EDA	Actividad Electro dérmica (Electrodermal Activity)
EEG	Electroencefalograma
EMG	Electromiograma
FGUI	Intefaz de usuario flexible (Flexible user interface)
GOMS-HRA	Goals, Operators, Methods, and Selection rules - Human Reliability Analysis
GSR	Respuesta Galvánica de la piel (Galvanic Skin Response)
HA	Actividad cardíaca (Heart Activity)
HCD	Diseño Centrado en las Personas (Human Centered Design)
HCI	Interacción Persona-Ordenador (Human-Computer Interaction)
HMI	Interfaz Persona-Máquina (Human-Machine Interface)
HR	Ritmo cardíaco (Heart Rate)
HRC	Colaboración Persona-Robot (Human Robot-Collaboration)
HRCAM	Modelo de Aceptación de la Colaboración Persona-Robot (Human-Robot Collaboration Acceptance Model)
HRI	Interacción Persona-Robot (Human-Robot Interaction)
HRIX	Experiencia de Interacción Persona-Robot (Human-Robot Interaction Experience)
HRV	Variabilidad del ritmo cardíaco (Heart Rate Variability)
HUBOXT	Human-Robot-Context
IIS	Puntuación de Impacto del Ítem (Item Impact Score)
IoT	Internet de las Cosas (Internet of Things)
ITPX	Experiencia Individual en el Desempeño de Tareas (Individual Task Performance eXperience)
LR	Lento-Rápido
MWL	Carga mental (Mental Workload)
PC	Confort Postural (Postural Confort)

PD	Diámetro Pupilar (Pupil Diametre)
PI	Pregunta de investigación
PPG	Fotopleletismografía (Photoplethysmography)
PWL	Carga postural de trabajo (Postural Workload)
RA	Realidad Aumentada
RL	Rápido-Lento
RPM	Respiraciones Por Minuto
RSI	Lesión por esfuerzo repetitivo (Repetitive Strain Injury)
RULA	Evaluación rápida de las extremidades superiores (Rapid Upper Limb Assessment)
RV	Realidad Virtual
SE-HRI	Autoeficacia – Interacción Persona-Robot (Self Efficacy - Human Robot Interaction)
SLR	Revisión Sistemática de la Literatura (Systematic Literature Review)
SORO	Operación de Robots Orientada al Servicio (Service Oriented Robot Operation)
ST	Temperatura cutánea (Skin Temperature)
TGUI	Interfaz de Usuario Tradicional (Traditional user interface)
TIC	Tecnologías de la Información y Comunicación
TL	Tarea Lenta
TR	Tarea Rápida
TRL	Nivel de Preparación Tecnológica (Technology Readiness Level)
TTF	Teoría de Ajuste Tarea-Tecnología (Task Technology Fit Theory)
UCAD	Diseño ágil centrado en el usuario (User Centered Agile Design)
UI	Interfaz de Usuario (User Interface)
UT	Tiempo del Usuario (User Time)
UX	Experiencia de Usuario (User Experience)
UXI	Indicador de Experiencia de Usuario (User Experience Indicator)
WOI	Workstation-Operator Interaction
WWL	Carga de trabajo ponderada (Weighted workload)
XC	eXperience Capturer

Capítulo I

Introducción

1 INTRODUCCIÓN

En este capítulo se aborda la introducción al tema. Se inicia con una explicación detallada del objeto de la tesis, seguida por una contextualización y motivación. Asimismo, se presentan las hipótesis planteadas y los objetivos específicos que se persiguen en esta tesis. Seguido, se describe la metodología empleada para llevar a cabo la investigación. Para finalizar, se expone el esquema de la tesis y se describen los capítulos restantes. El presente capítulo busca proporcionar una visión clara y completa del tema de la tesis, así como los objetivos, la metodología y la estructura general del trabajo.

1.1 Objeto de la tesis

La presente investigación se centra en una de las tecnologías más relevantes para la Industria 5.0, la Interacción Persona-Robot (HRI, por sus siglas en inglés). Concretamente en aquellos escenarios en los que personas y robots comparten el mismo espacio de trabajo para alcanzar un objetivo común de forma sincronizada, cooperativa o colaborativa. Se enfoca, por lo tanto, en la definición y validación de un modelo que permita evaluar la experiencia de usuario (UX, por sus siglas en inglés) y la aceptación tecnológica en HRI en entornos industriales.

A pesar del crecimiento en el uso de robots en la industria en los últimos años, aún existe una brecha en la comprensión de cómo las personas usuarias perciben y experimentan la interacción con estos robots en su desempeño laboral. La creación de un modelo sólido y validado para evaluar la UX en dichos entornos puede identificar áreas de mejora en la HRI, mejorando así la eficiencia y seguridad en la industria.

La investigación busca contribuir al avance del campo de la robótica en la industria y proporcionar un modelo útil para los desarrolladores de robots, diseñadores e investigadores del ámbito. Con los resultados de esta investigación, se espera proveer un mayor entendimiento de la UX en interacciones persona-robot en entornos industriales para así mejorar su bienestar, y a su vez, promover el desarrollo de robots más eficientes y seguros en la industria.

1.2 Contexto de la tesis

Esta sección tiene como objetivo contextualizar el tema de la tesis en relación con la continua evolución de la tecnología y su impacto en la sociedad, dando lugar a un nuevo paradigma socioeconómico. En este sentido, se exponen los aspectos más relevantes que están influyendo en la evolución hacia la nueva industria. En primer lugar, la situación de la industria hoy en día y en segundo lugar los antecedentes de este trabajo en el Centro de Innovación en Diseño (DBZ, por sus siglas en euskera) de Mondragon Unibertsitatea.

1.2.1 La industria hoy

La industria manufacturera está experimentando una revolución tecnológica, en la que la Industria 4.0 y 5.0 representan dos fases distintas (Coronado et al., 2022). La Industria 4.0 busca mejorar la eficiencia y la productividad a través de la fabricación inteligente, utilizando los pilares tecnológicos mencionados. En contraste, la Industria 5.0 se enfoca en una sociedad más sostenible y centrada en las personas, centrando su atención en la justicia social, el bienestar humano y la resiliencia de las industrias. En la Industria 4.0, el papel de las personas es limitado, mientras que en la Industria 5.0, la mano de obra humana es importante, respetando sus talentos, derechos, necesidades e identidades. Las tecnologías clave de la Industria 5.0

incluyen la colaboración persona-robot, los recursos renovables, la biónica, las tecnologías bioinspiradas y los materiales inteligentes (Coronado et al., 2022). La Industria 5.0 pretende crear escenarios altamente adaptables y personalizados en los que personas y robots puedan cooperar o colaborar para alcanzar objetivos comunes.

La industria 4.0 está “centrada en el rendimiento” (Coronado et al., 2022) y ha sido construida en torno a intereses económicos y tecnológicos, similar a las revoluciones industriales anteriores, que buscaban aumentar la productividad y lograr la producción en masa mediante nuevas tecnologías (Xu et al., 2021). Las revoluciones anteriores, como la Industria 1.0 impulsada por vapor, la Industria 2.0 impulsada por la electricidad y la Industria 3.0 impulsada por la electrónica y tecnologías de la información, han sido una importante fuente de crecimiento económico a lo largo de los años. Sin embargo, los problemas sociales y planetarios relacionados con las actividades industriales actuales están empezando a impulsar un cambio de paradigmas (Comisión Europea, 2021). Debido a ello, existe una escasa inclinación a aceptar y confiar en la tecnología de la automatización (Schwabe & Castellacci, 2020; Złotowski et al., 2017). Las personas trabajadoras poco cualificadas y de cualificación media, que realizan tareas rutinarias, tienen una tendencia a rechazar la tecnología de automatización, ya que pueden percibirla como una amenaza para sus empleos, su identidad, su singularidad y su seguridad (Złotowski et al., 2017).

La Industria 5.0 es un concepto reciente adoptado por la Comisión Europea (2021) que persigue la creación de industrias centradas en las personas, sostenibles y resilientes. A diferencia de las revoluciones industriales pasadas, que se centraban en la producción en masa y subestimaban los costos humanos y planetarios, la Industria 5.0 se enfoca en respetar la presencia, el talento y los derechos de las personas. Para lograr esto, propone la introducción de tecnologías que potencien el talento y la diversidad de las personas trabajadoras y salvaguarden los derechos humanos fundamentales. También se enfoca en la creación de procesos de producción sostenibles y en la reducción de residuos industriales, así como en la creación de industrias más ágiles, flexibles y adaptables (Comisión Europea, 2021).

La Tabla 1 resume las diferencias entre la Industria 4.0 y la Industria 5.0 (Comisión Europea, 2021; Demir et al., 2019; Nahavandi, 2019; Xu et al., 2021). Por lo tanto, mientras que la principal prioridad de la Industria 4.0 es la automatización de procesos, que intrínsecamente produce una reducción de la intervención humana en los procesos de fabricación, la Industria 5.0 puede devolver la fuerza humana a las fábricas y promover más empleos cualificados en comparación con la Industria 4.0.

Tabla 1: Diferencias entre la industria 4.0 y la industria 5.0 (Coronado et al., 2022).

Característica	Industria 4.0	Industria 5.0
Lema	Fabricación Inteligente.	Co-trabajo persona-robot y Bioeconomía.
Motivación	Alcanzar la producción en masa y aumentar los beneficios económicos.	Sociedad inteligente, justicia social, industrias resilientes, bienestar humano y sostenibilidad.
Rol humano	Las máquinas sustituyen a las personas.	Devolver la fuerza humana a las fábricas respetando los talentos, derechos, necesidades e identidades de las personas.
Tecnologías clave	Internet de las cosas, computación en nube, Big Data, robótica e inteligencia artificial.	HRC, recursos renovables, biónica, tecnologías bioinspiradas y materiales inteligentes.
Escenario típico en robótica	La interacción entre personas y máquinas/robots se limita a la programación y supervisión fuera de línea.	Escenarios altamente adaptables y personalizados en los que personas y robots pueden cooperar o colaborar para alcanzar objetivos comunes.

La introducción de las nuevas tecnologías plantea el reto de definir con mayor precisión el rol de las personas en un contexto en el que las revoluciones tecnológicas tendrán un impacto en los individuos y sus actividades, así como en la organización del trabajo (Pacaux-Lemoine et al., 2021).

En el contexto de la Industria 4.0, la persona operaria seguirá teniendo gran importancia, denominándose operador 4.0 (Romero et al., 2016). El concepto de Operador 4.0 se refiere al desarrollo del trabajo en fábricas hacia un enfoque de trabajo basado en el conocimiento, que es más exigente pero también más enriquecedor y flexible. Esta persona operaria, tiene como objetivo crear relaciones de confianza basadas en la interacción entre las personas y las máquinas haciendo posible que esas fábricas inteligentes aprovechen no sólo las fortalezas y capacidades de las máquinas inteligentes, sino que también doten a sus "operadores inteligentes" de nuevas habilidades para aprovechar las oportunidades que están siendo creadas por las tecnologías de la Industria 4.0 (Romero et al., 2016). La mano de obra humana, por lo tanto, asume un papel crucial dentro de las fábricas inteligentes (Hadorn et al., 2015). En este sentido, es necesario optimizar la interacción entre las máquinas y las personas a través de las nuevas interfaces, adaptándolas a sus necesidades para que así puedan emplear todo su conocimiento y habilidades explotando las posibilidades que ofrecen las nuevas tecnologías. Por ello, la UX, se podría considerar el eslabón que mayor vínculo genera entre el producto y la persona, y el lugar donde mayor presencia tienen las emociones humanas.

En este sentido, Pacaux-Lemoine et al. (2021) indican que, a nivel individual, la Industria 4.0 empleará menos personal no cualificado que la fábrica tradicional. Lo más probable es que las actividades rutinarias, caracterizadas por un bajo nivel de destreza manual o requisitos de interacción social, sean sustituidas por tecnologías (Fantini et al., 2020). Por otro lado, aumentará la necesidad de personal cualificado y bien formado. De este modo, las personas operarias manejarán información más abstracta, y a menudo estarán más "alejados" del proceso a controlar, con un mayor ámbito de supervisión (Pacaux-Lemoine et al., 2021). Se podría decir, que las personas operarias tendrán que tomar decisiones relevantes durante el proceso de manufactura para así poder afrontar situaciones inesperadas.

Además, Fantini et al. (2020) proponen una visión basada en el concepto de interacciones sociales. En esta perspectiva, una arquitectura multiagente de la fábrica social es propuesta por Romero et al. (2017). La persona operaria 4.0 está en el centro de una red social, formada por otros operadores sociales, pero también por máquinas y *software* calificados como "sociales". Esta arquitectura permite formalizar el diseño y la evaluación de las interacciones sociales del sistema de producción mediadas por interfaces.

Sin embargo, Colim et al. (2021) indican que son escasos los estudios empíricos que sobre la implementación de la robótica colaborativa que tienen en cuenta la ergonomía y los factores humanos. La colaboración persona-robot (HRC) se ha considerado una estrategia viable para ayudar a las personas trabajadoras asumiendo tareas peligrosas y/o físicamente exigentes (Colim et al., 2021). Esta relación mutua entre personas y robots da lugar a un potente marco de colaboración con un impacto positivo en la productividad, la flexibilidad y la creación de nuevos puestos de trabajo en lugar de sustituir a las personas trabajadoras (Villani et al., 2017).

En estos nuevos sistemas de trabajo, sin embargo, la seguridad y el bienestar de las personas siguen siendo retos. Según Maurice et al. (2017), es necesario seguir investigando en entornos industriales reales para validar la implementación de la HRC. Por lo tanto, es esencial desarrollar métodos de evaluación específicos y detallados para optimizar el diseño de estos puestos de trabajo con robótica colaborativa (Colim et al., 2020, 2021; Gualtieri, Palomba, et al., 2020; Maurice et al., 2017), teniendo en cuenta las condiciones industriales, las

características de las personas trabajadoras y los cobots, así como el nivel de interacción colaborativa (Khalid et al., 2017).

Aunque se anticipa la transición del Operador 4.0 al Operador 5.0, es importante reconocer que ambos paradigmas se encuentran actualmente en etapas de desarrollo (Nahavandi, 2019). Por otro lado, el concepto de la categoría "Operador 4.0 Saludable" surgió como parte del marco más amplio del Operador 4.0 (Romero et al., 2016) como una respuesta directa a las crecientes preocupaciones sobre los niveles crecientes de estrés entre la fuerza laboral y el estado general de bienestar psicológico y social (Buffet et al., 2013; Romero et al., 2018; Salanova et al., 2014) y empoderamiento (Sun et al., 2020). Implica la integración de datos de tecnologías portátiles, Internet de las cosas (IoT), inteligencia ambiental y técnicas de modelado para crear un gemelo digital de la persona operaria.

1.2.2 Interacción persona-robot

Los robots se presentan como uno de los nueve pilares de la industria 4.0 (Rüßmann et al., 2015). Sin embargo, las personas continuarán siendo de gran importancia para garantizar una alta flexibilidad en los procesos y para responder de manera proactiva a las necesidades cambiantes del mercado y a las solicitudes cada vez mayores de personalización de los productos (Prati et al., 2021). En este contexto, las fábricas inteligentes no solo necesitan robots, sino que también necesitan una fuerte Colaboración Persona-Robot (HRC). En estos escenarios, las tareas repetitivas, inseguras y físicamente exigentes suelen asignarse a los robots, mientras que las personas se encargarán del pensamiento crítico y la personalización (Maddikunta et al., 2022; Nahavandi, 2019).

En los últimos años, ha habido un aumento en la incorporación de robots en la industria 5.0 (Karabegović et al., 2020, 2022; Marvel et al., 2020), la cual tiene como pilar central el bienestar de las personas trabajadoras. Como resultado, las personas operarias deben interactuar con ellos diariamente mientras promueven su bienestar (Comisión Europea, 2021). Por lo tanto, es importante que se conviertan en socios en su vida diaria de manera simbiótica para que puedan tener una experiencia natural, fluida y satisfactoria (Boden et al., 2017; Chen et al., 2020; Kahn Jr et al., 2007; Lindblom & Alenljung, 2020). En los últimos años, se ha reconocido el potencial de las personas y los robots para trabajar juntos como enfoque viable para apoyar a las personas trabajadoras al asumir tareas peligrosas y físicamente exigentes (Colim et al., 2021). Esta relación simbiótica aprovecha las fortalezas respectivas de ambas partes para crear un marco de colaboración sólido que mejora la productividad, la flexibilidad y la generación de nuevas oportunidades de trabajo, en lugar de desplazar el trabajo humano (Villani et al., 2017). Los robots colaborativos son especialmente adecuados para abordar desafíos relacionados con tareas de fabricación y ensamblaje, ya que pueden interactuar físicamente con las personas en un espacio de trabajo compartido (Guerin et al., 2015; Matheson et al., 2019). Por otro lado, la intervención humana sigue siendo indispensable para garantizar un alto grado de adaptabilidad y capacidad de respuesta proactiva a las demandas constantemente cambiantes de la personalización del producto (Prati et al., 2021). Los sistemas robóticos poseen la capacidad de complementar y aumentar los atributos sensoriales, físicos y cognitivos humanos, mientras que las personas operarias pueden atender a las tareas cognitivas más complejas e intrincadas. Por lo tanto, el papel humano seguirá siendo relevante en la HRI.

Aunque el número de robots en la industria manufacturera ha aumentado de forma constante durante varias décadas, la tendencia actual de crecimiento es notable dado el número y la variedad de industrias que utilizan robots (Marvel et al., 2020). Tal y como indican Marvel et al. (2020), actualmente, la naturaleza "colaborativa" de los robots colaborativos es una mera cuestión de proximidad. Estas plataformas robóticas se utilizan en aplicaciones de automatización tradicionales y funcionan en gran medida de forma independiente del entorno

y los acontecimientos que los rodean. Aunque esto puede ser suficiente para las aplicaciones actuales, limita el potencial de la tecnología. A diferencia de la automatización estándar, los robots colaborativos (cobots) permiten una interacción más estrecha y segura entre las personas y las máquinas, aprovechando las ventajas de ambas partes, de acuerdo con las normas internacionales de funcionamiento y seguridad (ISO 10218, 2011; ISO/TS 15066, 2016).

Dautenhahn (2013) define la HRI como “la ciencia que estudia el comportamiento y las actitudes de las personas hacia los robots en relación con las características físicas, tecnológicas e interactivas de los mismos, con el objetivo de desarrollar robots que faciliten la generación de interacciones persona-robot que sean al mismo tiempo sean eficientes (de acuerdo con los requisitos originales de su área de uso prevista), aceptables para las personas, que satisfagan las necesidades sociales y emocionales de sus personas usuarias, y que respeten los valores humanos”.

La norma ISO 8373 (2021) define robot industrial como “manipulador multipropósito controlado automáticamente y reprogramable, programable en tres o más ejes, que puede estar fijo en su lugar o instalado en una plataforma móvil para su uso en aplicaciones de automatización en un entorno industrial”.

Para que los robots se conviertan en una parte natural del día a día de las personas usuarias es necesario que obtengan experiencias positivas a la vez de adecuadas de acuerdo con su propósito mediante una interacción fluida y satisfactoria (Boden et al., 2017; Chen et al., 2020; Kahn Jr et al., 2007; Lindblom & Alenljung, 2020). Esto significa que las personas deben experimentar que un robot cumple con los objetivos explícitos e implícitos existentes, lo que significa que actúa de manera eficiente y que consiga que las personas sientan confianza, seguridad y comodidad mientras conviven (Benyon, 2019).

Aunque las capacidades de los sistemas robóticos siguen ampliándose, muchas tareas en el ámbito industrial requieren un nivel de juicio, destreza y flexibilidad en la toma de decisiones que supera las capacidades de los robots actuales, lo que hace que las tareas pasen a ser dominadas por las personas (Lasota & Shah, 2015). Sin embargo, también hay muchas tareas sin valor añadido en estos ámbitos que podrían ser realizadas por asistentes robóticos. En este sentido, permitir que los robots colaboren con las personas en espacios de trabajo compartidos y realicen estas tareas tiene el potencial de aumentar la productividad y la eficiencia, lo que supone un fuerte incentivo para el desarrollo de tecnologías que fomenten esta colaboración (Lasota & Shah, 2015).

De acuerdo con Thrun (2004), las tecnologías robóticas de colaboración pueden agruparse en tres categorías principales: la robótica industrial, la robótica de servicios profesionales y la robótica de servicios personales. A diferencia de los manipuladores de robots, que son principalmente pasivos y no emplean sensores ni accionadores, la robótica colaborativa está constituida por dispositivos flexibles e “inteligentes” que permiten una interacción directa con la persona usuaria, creando una especie de “diálogo” con él para apoyar la actividad humana utilizando diferentes tipos de tecnologías (por ejemplo, sensores, accionadores y procesamiento de datos).

Para pasar del paradigma tradicional de la robótica de fabricación a una nueva generación de diseños fáciles de usar, Marvel et al. (2020) indican que deben cumplirse dos criterios:

- i) Demostrar que los diseños centrados en las personas son realmente más eficaces y/o eficientes. Si bien las funciones robóticas más avanzadas pueden estar menos orientadas a las personas usuarias ocasionales, se espera que los procesos básicos de programación, mantenimiento y producción se beneficien en términos de tiempo de aprendizaje, tiempo de ejecución y calidad de los resultados.

- ii) Comprobar que la respuesta de los operarios a la tecnología es positiva. La eficacia, eficiencia y la satisfacción de la persona usuaria constituyen la base de los principios de diseño centrado en las personas, tal y como se describe en la norma ISO 9241-210 (2019).

A fin de alcanzar este objetivo, se han llevado a cabo varias investigaciones durante los últimos años. Tal y como indican Lasota & Shah (2015), el primer paso hacia la creación de robots que colaboran de forma exitosa con personas es permitir que los robots se muevan por un espacio de trabajo compartido, lo que requiere el desarrollo de algoritmos especializados de planificación de trayectorias y marcos diseñados teniendo en cuenta el elemento humano.

1.2.3 La importancia de la Aceptación Tecnológica y la Experiencia de Usuario en la HRI

1.2.3.1 La Aceptación Tecnológica

El Modelo de Aceptación de la Tecnología (TAM, por sus siglas en inglés) es un marco teórico ampliamente utilizado que tiene como objetivo comprender y predecir la aceptación de la tecnología por parte de las personas usuarias (Davis, 1989, 1993). En el contexto de la HRI, el TAM puede ser aprovechado para explorar los factores que afectan la disposición de las personas usuarias a aceptar e interactuar con agentes robóticos. Al examinar la utilidad percibida y la facilidad de uso de los robots, así como las actitudes hacia ellos, el TAM ofrece un enfoque integral para comprender la dinámica de la HRI. Este modelo puede ofrecer ideas valiosas para el diseño e implementación de sistemas robóticos efectivos que se adapten mejor a las necesidades y expectativas de las personas usuarias.

El campo de la aceptación de la tecnología ha sido estudiado a través de varios modelos, comenzando con la teoría de difusión de Rogers en 1962, que propuso un modelo de cinco pasos desde la conciencia hasta la confirmación de la nueva tecnología (Bröhl et al., 2019; Rogers et al., 2014). Sin embargo, el TAM ha demostrado ser uno de los modelos más influyentes y ha sido utilizado en una amplia gama de contextos tecnológicos. Al centrarse en la percepción de la persona usuaria sobre la utilidad y facilidad de uso, el TAM proporciona una comprensión profunda de los procesos cognitivos y emocionales que influyen en la adopción de la tecnología.

El estudio de Bröhl et al. (2019) explora el modelo de aceptación de la colaboración persona-robot basado en los TAMs. El TAM es un marco teórico ampliamente utilizado que pretende comprender y predecir la aceptación de la tecnología por parte de las personas usuarias (Davis, 1989, 1993). En el contexto de la HRI, el TAM puede aprovecharse para explorar los factores que afectan a la disposición de las personas usuarias a aceptar e interactuar con robots. Al examinar la utilidad percibida y la facilidad de uso de los robots, así como las actitudes hacia ellos, el TAM proporciona un enfoque global para comprender la dinámica de la HRI. Este modelo puede ofrecer valiosas perspectivas para el diseño y la implantación de sistemas robóticos eficaces que se adapten mejor a las necesidades y expectativas de las personas usuarias. El campo de la aceptación tecnológica se ha estudiado a través de varios modelos, empezando por la teoría de la difusión de Rogers en 1962, que propuso un modelo de cinco pasos desde la concienciación hasta la confirmación de la nueva tecnología (Bröhl et al., 2019; Rogers et al., 2014). Sin embargo, los robots son percibidos y evaluados de forma diferente por las personas en comparación con otras tecnologías (Meissner et al., 2020). Como resultado de sus características distintivas, los modelos de aceptación de diferentes ámbitos pueden no ser aplicables en el campo de la HRI. Los robots poseen características únicas, como el movimiento autónomo y la interacción con su entorno, que los diferencian de otros dispositivos tecnológicos. Además, ofrecen formas de

comunicación innovadoras que hacen que las personas esperen modos de comunicación socialmente más hábiles e inteligentes que los que ofrecen otras tecnologías.

En este sentido, Bröhl et al. (2019) realizaron un estudio sobre la aceptación de la HRI en un contexto industrial y encontraron que la relevancia laboral era el predictor más significativo de la utilidad percibida, seguido de la norma subjetiva, la calidad de salida y el resultado en demostrabilidad en el modelo de aceptación. En cuanto a las variables de anclaje, las percepciones de control externo, autoeficacia y ansiedad ante los robots fueron las más relevantes. Por otro lado, el disfrute percibido, la seguridad percibida y la seguridad laboral fueron los mejores predictores de la facilidad de uso percibida, entre las variables de ajuste. Las implicaciones sociales, de protección de datos y éticas fueron comparativamente menos relevantes. El diseño ergonómico puede influir positivamente en las variables de ajuste y mejorar la facilidad de uso percibida y la intención de comportamiento de la HRI. Los resultados indican que el modelo original de aceptación de la tecnología puede aplicarse al dominio de la HRI, con altos coeficientes de correlación entre la utilidad percibida, la facilidad de uso percibida, la intención conductual y el comportamiento de la persona usuaria (Bröhl et al., 2019).

La Figura 1 detalla específicamente en qué consiste el Modelo de Aceptación de la Colaboración Persona-Robot (HRCAM, por sus siglas en inglés). En cuanto a la **utilidad percibida** el predictor más importante en el HRCAM es la importancia del trabajo, seguido de la norma subjetiva, que podría deberse al contexto laboral e industrial del robot. La **facilidad de uso percibida** está influida por las variables de anclaje (que incluyen la autoeficacia, la ansiedad del robot, la afinidad de la tecnología, la percepción de control externo y la edad) y el ajuste (que incluyen el disfrute percibido, implicaciones sociales, implicaciones legales, que a su vez incluye la seguridad laboral y la protección de datos, las implicaciones éticas y la seguridad percibida). Según En cuanto a las variables de anclaje, los coeficientes de correlación más altos se encontraron para las variables del modelo tradicional TAM 3 autoeficacia, ansiedad ante el robot y percepción de control externo.

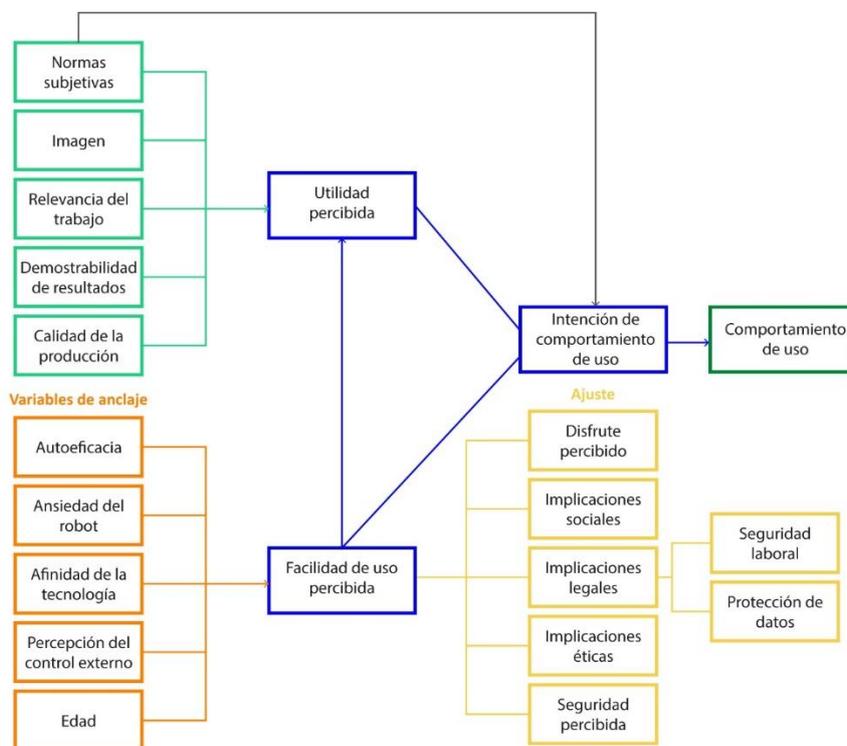


Figura 1: El modelo HRCAM. Adaptado de Bröhl et al. (2019).

1.2.3.2 La Experiencia de Usuario, Usabilidad, Ergonomía y Hedonomía en HRI

Tal y como indican Coronado et al. (2022), los factores de calidad establecidos en la norma ISO 9241-11 (2018) presentan un importante solapamiento y diferentes conceptualizaciones. Los dos conceptos que presentan un mayor solapamiento son la usabilidad y la UX. Por un lado, la usabilidad, en algunos casos, se relaciona con la "facilidad de uso". Sin embargo, su concepto es más amplio. Por un lado, la norma ISO 9241-210 (2019) define la Experiencia de Usuario (UX) como "las percepciones y respuestas de una persona que resultan del uso o el uso anticipado de un producto, sistema o servicio", incluidas las emociones, creencias, preferencias, percepciones, respuestas físicas y psicológicas, comportamientos y logros de las personas usuarias que se producen antes, durante y después del uso. Por otro lado, la norma ISO 9241-11 (2018), define la usabilidad como "la medida en que un sistema, producto o servicio puede ser utilizado por personas usuarias específicas para lograr metas específicas con eficacia, eficiencia y satisfacción en un contexto de uso específico".

Sin embargo, se pueden identificar dos elementos diferentes: aquellos relacionados con factores objetivos y orientados al rendimiento (eficacia y eficiencia) y aquellos relacionados con aspectos subjetivos (satisfacción) (Diefenbach et al., 2014). A pesar de esta definición estandarizada, parece no existir un consenso en las comunidades de HCI e HRI sobre la definición de usabilidad (Rajanen et al., 2017). Por lo tanto, varios autores proponen diferentes atributos que componen la definición de usabilidad.

Ejemplos de artículos de revisión que resumen las diferentes definiciones de usabilidad son (Coronado et al., 2022; Rajanen et al., 2017; Rana & Kumar Dubey, 2010). Según Lachner et al. (2016), algunos autores consideran la UX como un subconjunto del componente de satisfacción de la usabilidad. En contraste, otros pueden considerar la usabilidad como un subconjunto de UX. Además, una tercera perspectiva considera que la usabilidad enfatiza las medidas objetivas y la UX enfatiza las medidas subjetivas. Para reducir la confusión entre los conceptos de usabilidad y UX, Sauer et al. (2020) propusieron un modelo holístico diseñado para ser coherente con las definiciones de las normas ISO. Este modelo integra el enfoque holístico de la UX y considera tanto elementos subjetivos como objetivos. Además, elementos relacionados con las emociones, como el placer, la aceptación, la confianza y la estética, se consideran fuera del ámbito de la usabilidad, lo cual es un enfoque aceptado por muchos profesionales en muchos casos.

Teniendo en cuenta los diferentes enfoques, Coronado et al. (2022) proponen un modelo conceptual holístico para la HRI. La Figura 2 muestra la relación y los límites existentes entre la usabilidad, UX, hedonomía y ergonomía.

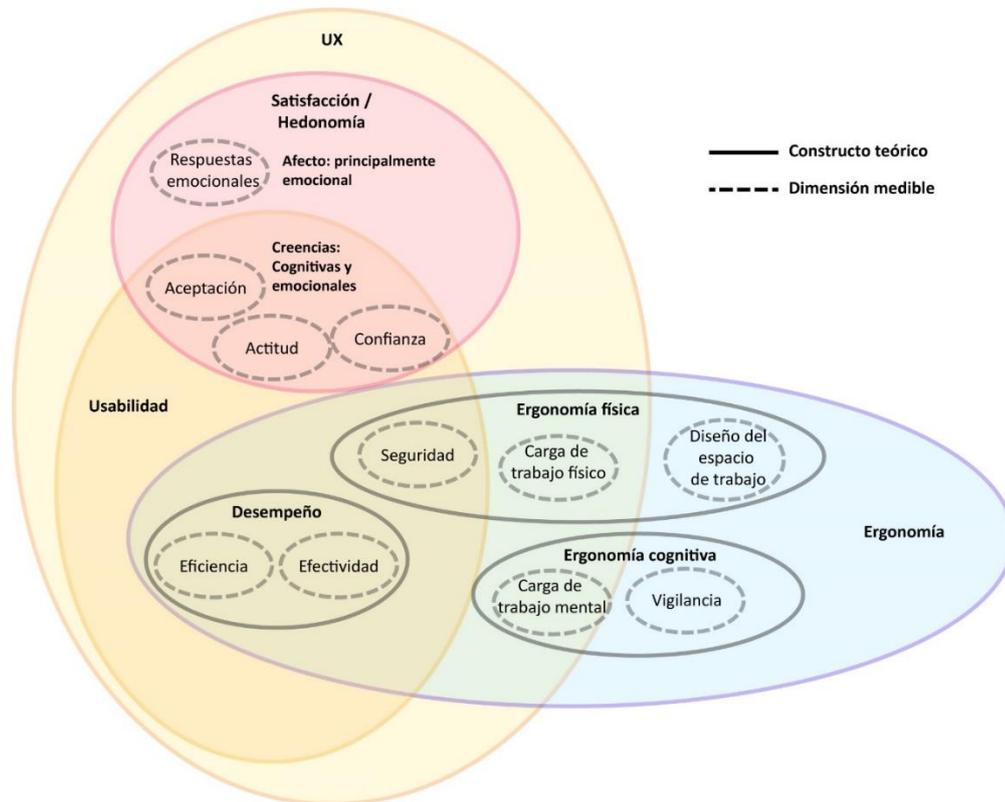


Figura 2: Factores de calidad más representativos analizados en la literatura sobre HRI. Adaptado de Coronado et al. (2022).

Para comenzar, Coronado et al. (2022) agrupan los **factores hedonómicos**. La hedonomía representa una continuación conceptual de la ergonomía centrado en "los aspectos agradables o placenteros de la interacción entre la persona y la tecnología" (Hancock et al., 2005). Según Oron-Gilad & Hancock (2017), el fundamento moral o núcleo principal de la ergonomía se enfoca en reducir el dolor, las lesiones y el sufrimiento en el lugar de trabajo. Sin embargo, esta disciplina a menudo se limita a mostrar la importancia de prevenir eventos negativos que "eventualmente no suceden" (Oron-Gilad & Hancock, 2017). Por el contrario, la hedonomía se centra en aspectos más positivos de las interacciones laborales al "promover la aparición de interacciones satisfactorias, que se pueden comprobar u observar" (Oron-Gilad & Hancock, 2017). Este paradigma centrado en la satisfacción y lo afectivo propuesto en las disciplinas de la hedonomía contrastan con el enfoque predominantemente orientado a la seguridad y la productividad de la investigación tradicional en ergonomía. Hancock et al. (2005) proponen la pirámide Hedonómica de Hancock (Figura 3), que se basa en la jerarquía de necesidades psicológicas de Maslow (1943), aclara los límites tanto de la hedonomía como de la ergonomía. Esta pirámide comienza en la parte inferior definiendo aspectos que pueden cumplir objetivos colectivos y funcionales. Cada nivel superior de la pirámide se centra cada vez más en aspectos individuales y no funcionales. Además, los factores de usabilidad se dividen en aquellos más cercanos a la definición de hedonomía (principalmente subjetivos) y aquellos presentados tradicionalmente en la ergonomía (principalmente objetivos).



Figura 3: Pirámide hedonómica de Hancock. Adaptado de Coronado et al. (2022).

En este sentido, Coronado et al. (2022) agrupan como factores hedonómicos aquellos que promueven la creación de experiencias interactivas capaces de mantener niveles emocionales óptimos es importante para reducir los niveles de estrés, evitar errores y mejorar el rendimiento de las tareas. Además, los factores relacionados con lo hedónico, como la felicidad, la estabilidad emocional y las emociones positivas, suelen considerarse dimensiones relevantes que reflejan el bienestar de las personas. El modelo conceptual propuesto por Coronado et al (2022) en la Figura 2 clasifica los factores hedónicos en dos grupos. El primer grupo considera aquellos factores predominantemente influenciados por aspectos emocionales. En este grupo, el concepto de mayor nivel es el afecto, que a menudo se utiliza para incluir términos relacionados con las emociones. Por otro lado, el segundo grupo considera aquellos factores en los que tanto los aspectos emocionales como los cognitivos son relevantes. A diferencia de los factores puramente emocionales, las creencias están recibiendo más atención en la comunidad de HRI con enfoque industrial, siendo la confianza el aspecto hedónico más común evaluado o discutido.

La norma ISO 6385 (2016) establece la ergonomía como una "disciplina científica preocupada por comprender las interacciones entre las personas y otros elementos de un sistema, y la profesión que aplica teoría, principios, datos y métodos para el diseño con el fin de optimizar el bienestar humano y el rendimiento general del sistema". Sin embargo, la concepción más común de la ergonomía se refiere a cómo las empresas diseñan tareas, escenarios e interfaces capaces de maximizar la eficiencia y las condiciones de trabajo de sus empleados (Oron-Gilad & Hancock, 2017). La mayoría de los trabajos en la literatura identifican dos áreas principales de la ergonomía: la ergonomía física y la ergonomía cognitiva. La **ergonomía** se centra comúnmente en dos objetivos principales. El primer objetivo es optimizar el bienestar mental y físico de las personas al prevenir el dolor y las situaciones de riesgo al interactuar o trabajar con máquinas. El segundo objetivo es optimizar el rendimiento del sistema al mejorar su usabilidad objetiva y funcionalidad. Se dividen por lo tanto los factores de ergonomía en tres clases principales: rendimiento, ergonomía física y ergonomía cognitiva. La ergonomía física y la ergonomía cognitiva son las clasificaciones más utilizadas en ergonomía. Por un lado, la ergonomía física se ocupa de los posibles efectos negativos o consecuencias en el cuerpo humano producidos por situaciones de trabajo, como posturas, trabajo pesado, movimientos repetitivos o fuerzas. En este contexto, el objetivo principal es construir sistemas interactivos y entornos de trabajo que sean compatibles con el tamaño, la fuerza y las capacidades físicas de las personas usuarias, y que al mismo tiempo no creen riesgos adicionales para la salud o lesiones. Por otro lado, los factores de la ergonomía cognitiva se centran en la creación de sistemas que se adapten a las capacidades perceptivas y

psicológicas de las personas usuarias, lo que les permite comprender el estado del entorno y razonar al respecto. A diferencia de los factores hedonómicos, donde las emociones pueden tener una influencia considerable, esta clase incluye aquellos factores en los que principalmente se requieren capacidades cognitivas y racionales, y donde los elementos cognitivos y perceptuales pueden verse potencialmente influenciados de manera negativa. Otra diferencia es que las creencias y los factores afectivos se pueden medir, cambiar o influir antes, durante y después de la interacción con los robots, mientras que los factores incluidos en la ergonomía cognitiva y la ergonomía física se miden predominantemente o son relevantes durante la interacción con los robots.

1.2.3.3 La UX en el tiempo

De acuerdo con Roto et al. (2011), el núcleo de la UX radica en la experiencia real de uso, sin embargo, esto no cubre todas las preocupaciones relevantes de UX. Las personas pueden tener experiencias indirectas antes de su primer encuentro a través de las expectativas formadas a partir de la experiencia existente en tecnologías relacionadas, presentaciones, demostraciones u opiniones de otros. De manera similar, la experiencia indirecta se extiende después del uso, por ejemplo, a través de la reflexión sobre el uso anterior o mediante cambios en las valoraciones de uso de las personas.

Además, las contrastantes entre "experimentar" y "una experiencia", plantean la pregunta del lapso de tiempo apropiado al centrarse en la UX (Roto et al., 2011). En este sentido, tal y como indican (Roto et al., 2011), la UX puede referirse a un cambio específico de sentimiento durante la interacción (UX Momentánea), la evaluación de un episodio de uso específico (UX Episódica) o las opiniones sobre un sistema en su totalidad después de haberlo utilizado durante un tiempo (UX Acumulada). La UX Anticipada puede relacionarse con el período antes del primer uso o cualquier otro de los tres lapsos de tiempo de UX, ya que una persona puede imaginar un momento específico durante la interacción, un episodio de uso o la vida después de utilizar un sistema (Figura 4).

Centrarse en el momento puede proporcionar información sobre las respuestas emocionales de una persona a los detalles de la interfaz de usuario (UI, por sus siglas en inglés). Centrarse en períodos más largos puede revelar el impacto eventual de las experiencias momentáneas en la UX acumulada. Por ejemplo, la importancia de una fuerte reacción negativa durante el uso puede disminuir después de resultados exitosos, y la reacción puede ser recordada de manera diferente. Un enfoque en la experiencia momentánea plantea demandas diferentes en el diseño y la evaluación que un enfoque en los episodios de uso o lapsos de tiempo más largos.

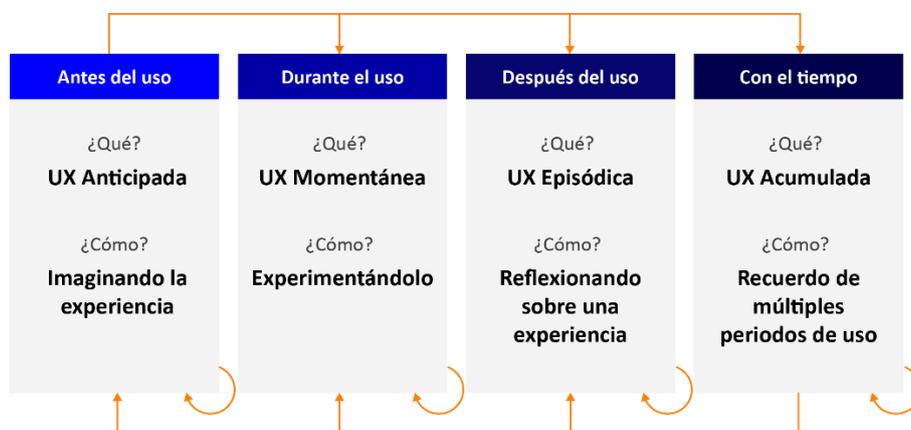


Figura 4: Los intervalos temporales de la experiencia de la persona usuaria, los términos para describir el tipo de UX relacionada con los intervalos, y el proceso interno que tiene lugar en los distintos periodos. Adaptado de Roto et al. (2011).

La UX es un término que se ha establecido en la investigación y la práctica de HCI. La UX denota que la interacción con un sistema tecnológico contemporáneo va más allá de la usabilidad y se extiende a las emociones antes, durante y después de usar el sistema, y no puede definirse únicamente estudiando los atributos fundamentales de la usabilidad, como la eficacia, eficiencia y la satisfacción de la persona usuaria. Medir la UX se convierte en una tarea más complicada cuando el objetivo de la interacción no es solo un sistema tecnológico o una aplicación, sino todo un entorno.

1.2.4 Antecedentes en el DBZ

En el marco de la industria 4.0, la aceptación tecnológica y la UX, el DBZ está trabajando para generar herramientas que contribuyan al desarrollo de la interacción persona-máquina. Algunos proyectos previos a este trabajo se han desarrollado con el objetivo de ayudar al conjunto de actores del binomio universidad-empresa, y se exponen como antecedentes de la presente investigación:

1.2.4.1 Metodología UCAD

Ante la carencia de procedimientos estructurados centrados en las personas usuarias para diseñar las Human Machine Interfaces (HMI), Aranburu Zabalo et al. (2017) crearon la metodología User Centered Agile Design (UCAD). Esta metodología nace de la unión de las metodologías Diseño Centrado en el Usuario (HCD) y Desarrollo de *Software* Ágil (DSA). El propósito de UCAD es proporcionar un método de trabajo para crear un *software* óptimo en términos de función, basado en las motivaciones y necesidades de la persona usuaria, y conseguir un producto amigable e intuitivo. Se trata de un proceso estructurado que muestra las fases, las herramientas y el modo de empleo para su aplicación en los sistemas digitales de los entornos de trabajo de la Industria 4.0. UCAD se compone de tres fases principales: Comprender, Diseñar y Desarrollar (Figura 5).

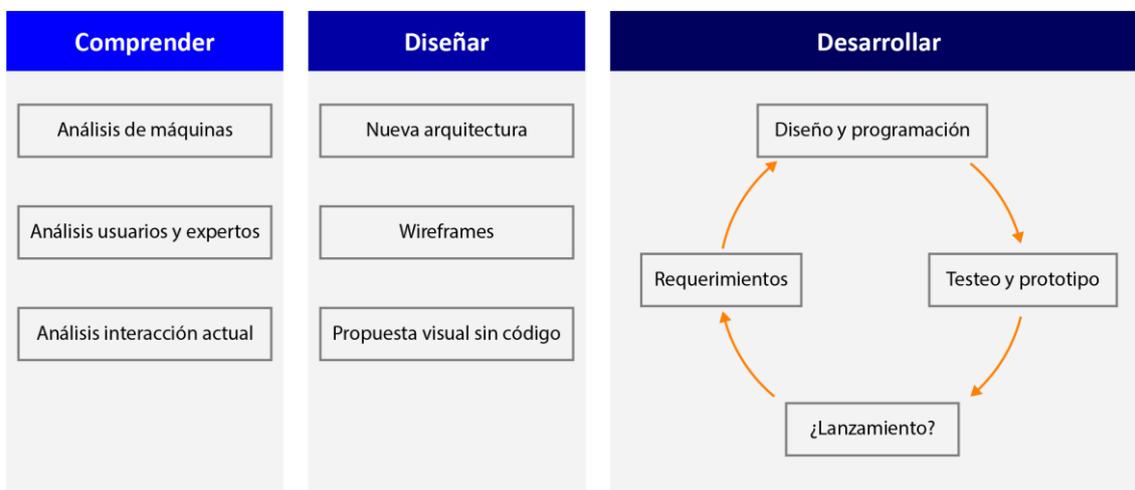


Figura 5: Metodología UCAD. Adaptado de Aranburu Zabalo et al. (2017).

1.2.4.2 Modelo DEMOX

Aranburu Zabalo (2020) ha desarrollado el modelo Design and Evaluation for Machine-Operator eXperiences (DEMOX), un modelo que permite evaluar la experiencia actual de los operarios con las interfaces industriales y que habilita la propuesta de nuevas soluciones (Figura 6).

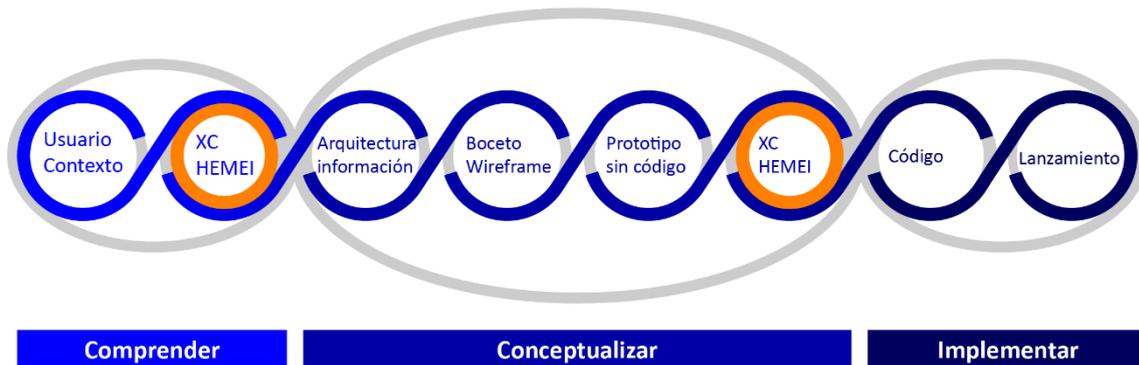


Figura 6: Metodología DEMOX. Adaptado de Aranburu Zabalo (2020).

- **Evaluación de usuario: eXperience Capturer (XC)**

Ante la necesidad de crear herramientas multimétodo que evalúen aspectos emocionales y de usabilidad en las tres fases de la evaluación, Aranburu et al. (2020) han desarrollado eXperience Capturer (XC). Su objetivo es analizar la experiencia de los operarios durante la interacción con interfaces industriales. Se trata de un proceso de evaluación compuesto por tres fases principales referidas a la interacción: previa, durante y posterior (Figura 7).

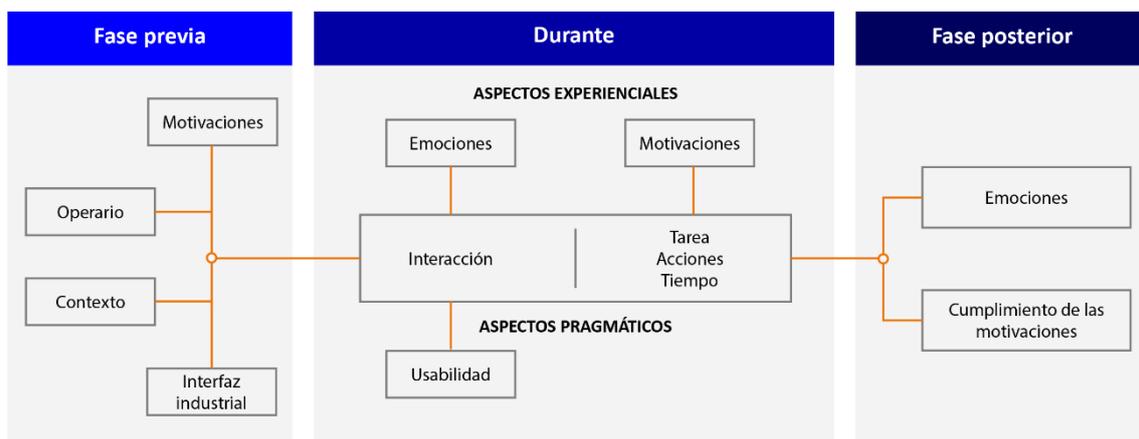


Figura 7: Herramienta XC. Adaptado de Aranburu et al. (2020).

La fase previa consiste en recoger datos sobre la persona usuaria, el contexto y la interfaz y definir las motivaciones. En la segunda fase (durante), se monitoriza la actividad de la persona usuaria mientras ejecuta una serie de tareas definidas por el equipo de evaluación, para poder detectar factores críticos de usabilidad y las respuestas emocionales. La tercera fase (posterior) consiste en evaluar las emociones y el cumplimiento de las motivaciones.

- **Evaluación experta: HEMEI**

HEMEI ofrece la valoración del carácter intencionado del producto sobre los aspectos pragmáticos y experienciales, mediante un proceso ágil que exige pocos recursos y permite evaluar la aplicación en su totalidad (Aranburu et al., 2019). Se basa en la definición de una serie de aspectos o principios para la evaluación de las interfaces industriales desde la perspectiva experta. Permite evaluar las interfaces tomando en cuenta tanto los aspectos pragmáticos como los experienciales de la interacción. Consta de 199 aspectos de evaluación, de los cuales 134 son de carácter pragmático y 65 experienciales. Los aspectos de evaluación se han dividido en 6 categorías, en base a los aspectos para cada tipo de aplicación. Estas son las categorías definidas: (i) General, (ii) Planning, (iii) Setup, (iv) Operation, (v) Monitoring y (vi)

Support and Maintenance. Mediante esta clasificación en categorías, se agiliza el proceso de evaluación de la herramienta, para que el evaluador solo tenga que centrarse en los aspectos de la aplicación en cuestión.

1.2.4.3 Modelo ITPX

Con el objetivo de comprender el impacto de los avances tecnológicos dentro de las empresas manufactureras, Mazmela Etxabe (2020) ha desarrollado el modelo *Individual Task Performance eXperience* (ITPX). Se trata de un modelo que proporciona un enfoque holístico y considera los aspectos relacionados con la UX analizando su impacto en el proceso de aprendizaje y toma de decisiones durante la ejecución de tareas (Figura 8). El modelo ITPX evalúa el proceso de ejecución de tareas en una interfaz industrial persona-máquina (HMI) desde tres perspectivas:

- i) La evaluación del valor de la TTF, que se lleva a cabo por un evaluador experto.
- ii) La evaluación de la UX en términos de intención de uso continuado y satisfacción personal, que son valores que se recogen de un cuestionario que lo cumplimentan las personas usuarias.
- iii) La evaluación de la ejecución de la tarea utilizando variables objetivas medidas por factores de tiempo y la tasa de error durante la ejecución de la tarea, las cuales se recogen durante la observación por un evaluador experto.

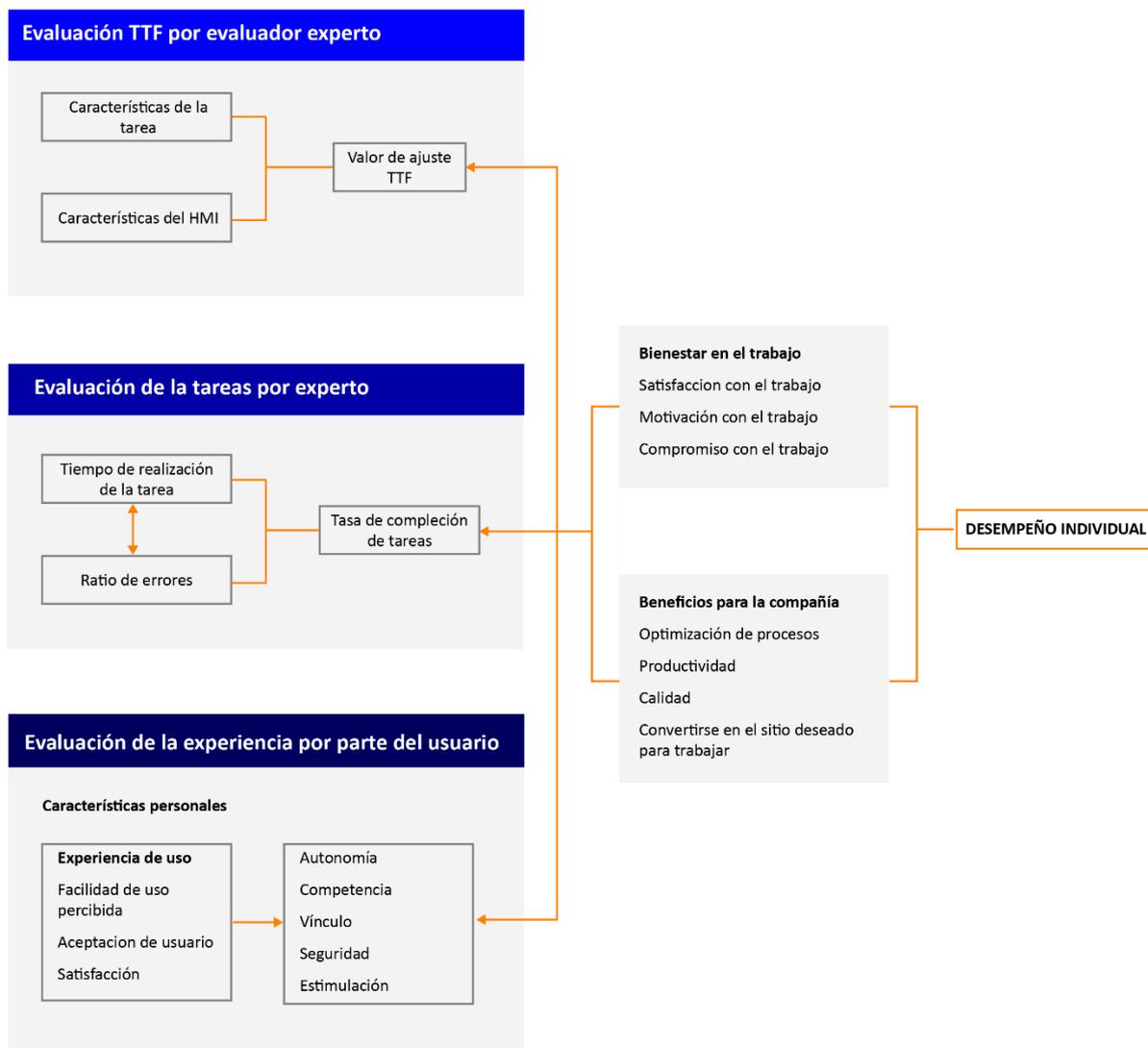


Figura 8: Modelo ITPX. Adaptado de Mazmela Etxabe (2020)

1.2.5 Interés del tema

El concepto de la HRI abarca diferentes niveles de interacción en los que los sistemas robóticos y las personas coexisten en espacios compartidos. Estos niveles van desde espacios de trabajo independientes hasta cooperación y colaboración. Si bien esta colaboración va más allá de la robótica social, desempeña un papel fundamental en la creciente complejidad y dinamismo de los procesos de fabricación modernos. A medida que estas interacciones trascienden lo puramente físico, el campo multidisciplinario de HRI se ha expandido para incluir estudios sociales.

El diseño juega un papel fundamental en el avance de las instituciones, corporaciones, empresas y, en su conjunto, de la cultura y la sociedad a nivel global. Es capaz de crear, innovar y proporcionar soluciones transversales, sostenibles, sociales, inclusivas y de alto impacto económico. En este sentido, el diseño se ha convertido en un elemento clave de sectores específicos como la fabricación avanzada y especialmente las relaciones persona-robot, como se refleja en el nuevo “Plan de Ciencia, Tecnología e Innovación Euskadi 2030” (PCTI 2030) y la estrategia “RIS3 Euskadi”. el “Plan de Recuperación, Transformación y Resiliencia” del Gobierno de España, el plan de recuperación para Europa “NextGenerationEU” y la estrategia Horizon Europe promovida por la Comisión Europea. Esta transformación debe situar a las personas en el centro de los procesos, para garantizar un futuro digital, verde e inclusivo para nuestra civilización y las venideras.

En la misma línea, la Asociación Europea de Investigación de Fábricas del Futuro (EFRA, 2023) destaca la importancia de comprender la transformación en el trabajo y la organización de las fábricas en la actualidad. Es crucial acelerar la adopción de los avances tecnológicos y promover un cambio significativo en la forma en que se lleva a cabo el trabajo en las fábricas. Es fundamental identificar cómo la tecnología ha impactado el trabajo, los perfiles laborales asociados y si ha contribuido a hacer que los lugares de trabajo promuevan el bienestar de las personas y sean más inclusivos. La necesidad de una mejora continua en el bienestar de las personas trabajadoras de fábrica y la atractividad de los lugares de trabajo se ve reforzada por la evolución demográfica. Los desarrollos centrados en las personas también deben contribuir a aumentar la capacidad general de producción de las fábricas, mejorar la eficiencia energética y el uso de recursos, entre otros aspectos. Para lograr esto, la tecnología debe desempeñar un papel fundamental, reduciendo los esfuerzos necesarios para la cooperación con las máquinas y disminuyendo el estrés y la tensión en las personas trabajadoras.

Además, el crecimiento significativo en la cantidad de revistas y special issues dedicados a investigaciones sobre la HRI en editoriales prominentes como Elsevier, Springer y Frontiers refleja la creciente importancia y el interés sostenido en este campo multidisciplinario.

En esta línea, la industria robótica está experimentando un crecimiento significativo, en el cual las personas y los robots serán capaces de colaborar (Harriott et al., 2013). Sin embargo, según Harriott et al. (2013), todavía no existe un modelo universal que evalúe el impacto de esta interacción en las personas y contemple la perspectiva de género. La inclusión de la perspectiva de género beneficia a la investigación y la innovación al mejorar la calidad y la pertinencia de la misma. Además, de acuerdo con el Plan de Igualdad de Género (PIG) del Programa Marco de Investigación e Innovación Horizonte Europa 2021-2027 (*Horizon Europe guidance on gender equality plans - Publications Office of the EU, s. f.*), las investigaciones deben contar con un plan de igualdad de género o una estrategia equivalente para poder optar a la financiación.

La realización de la presente tesis doctoral supondrá ser un referente en el campo de la HRI desde la perspectiva de la UX en entornos industriales. La Industria 4.0 está conduciendo a un nivel completamente nuevo de la automatización de los procesos, promoviendo una

interacción más estrecha entre persona y máquina y fomentando el uso eficiente tanto de las máquinas como de las personas en un entorno cooperativo. La Industria 5.0 redefine el papel de las personas y transforma los actuales puestos de trabajo hacia formatos aún desconocidos, y tiene como pilar fundamental el bienestar de las personas. En este contexto, las personas operarias seguirán siendo de gran importancia, por lo que resulta crucial optimizar la interacción entre las personas y las máquinas en los espacios colaborativos de trabajo. En consecuencia, la UX en estos espacios colaborativos se convierte en un aspecto clave en la relación entre las máquinas/robots y las personas. Trabajar en los aspectos de interacción tendrá un impacto notable en la optimización de los entornos de trabajo, mejorando las condiciones laborales y aumentando la efectividad en su uso.

Gracias a estas herramientas, será posible evaluar entornos de trabajo con el fin de proporcionar experiencias positivas, logrando generar motivación, implicación, eficiencia, interacciones satisfactorias y garantizar procesos de aprendizaje en sistemas tan complejos como los que se pueden dar en el sector industrial, un sector clave de la Comunidad Autónoma Vasca. Por lo tanto, esta investigación supondría una aportación de gran interés social, ya que al crear entornos de trabajo que proporcionen experiencias positivas no solo aumentan la eficiencia del sistema, sino que logran optimizar los entornos de trabajo, mejorando la calidad de trabajo de los operarios en su día a día. Además, contribuirá a poner el diseño al servicio de la transformación y recuperación económica de nuestro territorio y a que este se consolide, de manera definitiva, como la potente herramienta transversal de innovación y transformación que es. Concretamente, en el marco de la fabricación avanzada y los nuevos entornos automatizados donde las personas comparten espacios de trabajo con robots.

1.3 Hipótesis de la tesis

En esta sección se recogen las hipótesis planteadas para esta tesis doctoral, que son:

Hipótesis 1: En el contexto de HRI en un entorno industrial, cuanto mayor sea el indicador HEUROBOX incrementará los valores de valencia emocional y *engagement*, mientras que los valores de memorización y carga mental se verán reducidos.

Hipótesis 2: En el contexto de HRI en un entorno industrial, cuanto más elevados sean los valores de velocidad y aceleración del robot, incrementará la carga mental, mientras que los valores de la valencia emocional y la percepción general de la experiencia de uso serán disminuidos.

Hipótesis 3: En el contexto de HRI en un entorno industrial, el uso de interfaces multimodales conlleva una disminución en los valores de memorización y carga mental y un incremento de la experiencia de uso percibida en comparación con el uso de una interfaz unimodal.

1.4 Objetivos de la tesis

Para alcanzar los objetivos de la investigación, es necesario formular un objetivo principal cohesivo que pueda encapsular los retos subyacentes. Así pues, el objetivo principal de esta tesis doctoral es la siguiente:

Objetivo principal

Diseñar, definir y validar un nuevo modelo de evaluación centrado en la UX para la optimización de la interacción persona-robot en entornos industriales.

Este modelo permitirá identificar cuáles son los factores humanos que impactan durante estos procesos de fabricación y optimizarlos desde una visión holística. Así, el modelo deberá integrar la evaluación por parte de la persona experta, la evaluación basada en el test de usuario por tareas y la evaluación subjetiva sobre la percepción de usuario. Asimismo, se perseguirán los siguientes sub-objetivos relacionados con cada una de las hipótesis planteadas:

Siendo la primera hipótesis la siguiente **“Hipótesis 1: En el contexto de HRI en un entorno industrial, cuanto mayor sea el indicador HEUROBOX incrementará los valores de valencia emocional y engagement, mientras que los valores de memorización y carga mental se verán reducidos.”** se plantea el siguiente sub-objetivo:

Sub-objetivo 1: Desarrollar un nuevo modelo evaluación para investigar la interacción entre personas y robots en entornos industriales, con el fin de correlacionar el indicador HEUROBOX con la valencia emocional, el *engagement*, la memorización y la carga mental de las personas usuarias durante dicha interacción.

Siendo la segunda hipótesis la siguiente **“Hipótesis 2: En el contexto de HRI en un entorno industrial, cuanto más elevados sean los valores de velocidad y aceleración del robot, incrementará la carga mental, mientras que los valores de la valencia emocional y la percepción general de la experiencia de uso serán disminuidos.”**, se plantea el siguiente objetivo y subobjetivos:

Sub-objetivo 2: Validar el nuevo modelo a través de un caso de estudio que examine las respuestas fisiológicas (específicamente carga mental y valencia emocional) y la percepción de la persona usuaria en respuesta a distintas variables cinemáticas, específicamente, comparando la interacción con una velocidad y aceleración lenta versus una velocidad y aceleración rápida. Además, se busca identificar posibles correlaciones entre estos aspectos.

Siendo la tercera hipótesis la siguiente **“Hipótesis 3: En el contexto de HRI en un entorno industrial, el uso de interfaces multimodales conlleva una disminución en los valores de memorización y carga mental y un incremento de la experiencia de uso percibida en comparación con el uso de una interfaz unimodal.”**:

Sub-objetivo 3: Validar el nuevo modelo a través de un caso de estudio que examine las respuestas fisiológicas (específicamente la carga mental y la memorización) y la percepción de la persona usuaria en respuesta a diferentes interfaces, específicamente comparando una interfaz unimodal (solo visual) con una interfaz multimodal (visual + voz). Además, se busca identificar posibles correlaciones entre estos aspectos.

1.5 Metodología de la investigación

La presente investigación se enmarca como un estudio de tipo de tipo explicativo-aplicado. La metodología utilizada se ha estructurado en tres bloques principales:

- A- Enmarque científico-tecnológico:** Se llevará a cabo un estudio teórico exhaustivo que sustentará la estructura de la investigación. Este estudio comprenderá los siguientes aspectos: (i) análisis de los marcos teóricos y modelos existentes que conformarán la base para el diseño del nuevo modelo, (ii) examen de las medidas y métodos de medición utilizados en cada constructo, y (iii) revisión de los casos de estudio y pruebas de usabilidad existentes en la literatura.
- B- Desarrollo metodológico:** En esta etapa se desarrollará el nuevo modelo, el cual será definido mediante una serie de variables obtenidas del estudio teórico y se representará mediante relaciones conceptuales entre los constructos. Asimismo, se diseñarán las herramientas necesarias para realizar de manera óptima todas las mediciones.
- C- Propuesta y validación de la metodología:** Se realizará, por un lado, una validación parcial mediante expertos y una validación empírica mediante dos casos de estudio. Estos casos de estudio permitirán poner a prueba el modelo y verificar si las hipótesis planteadas son respaldadas por los resultados obtenidos.

Para llevar a cabo esta investigación, se ha empleado la metodología Design Inclusive Research (DIR) de Horváth (2007). Esta metodología tiene como objetivo obtener conocimientos para el diseño y la solución de problemas complejos mediante un enfoque multidisciplinar. El DIR incluye una fase de acciones creativas que se sitúa entre las etapas de exploración y evaluación de la investigación. Su finalidad es generar conceptos, modelos y metodologías, probar su factibilidad y viabilidad mediante iniciativas de prueba y experimentación, y lograr una mejor comprensión y perfeccionamiento.

El DIR se ha considerado la metodología adecuada para esta investigación, ya que las experimentaciones son fundamentales para validar las hipótesis planteadas. La Figura 9 muestra los dos casos de estudios enmarcados en esta investigación que se han llevado a cabo a lo largo del proyecto. Además, se ha realizado una vigilancia continua del estado del arte dado el creciente número de investigaciones en esta disciplina.

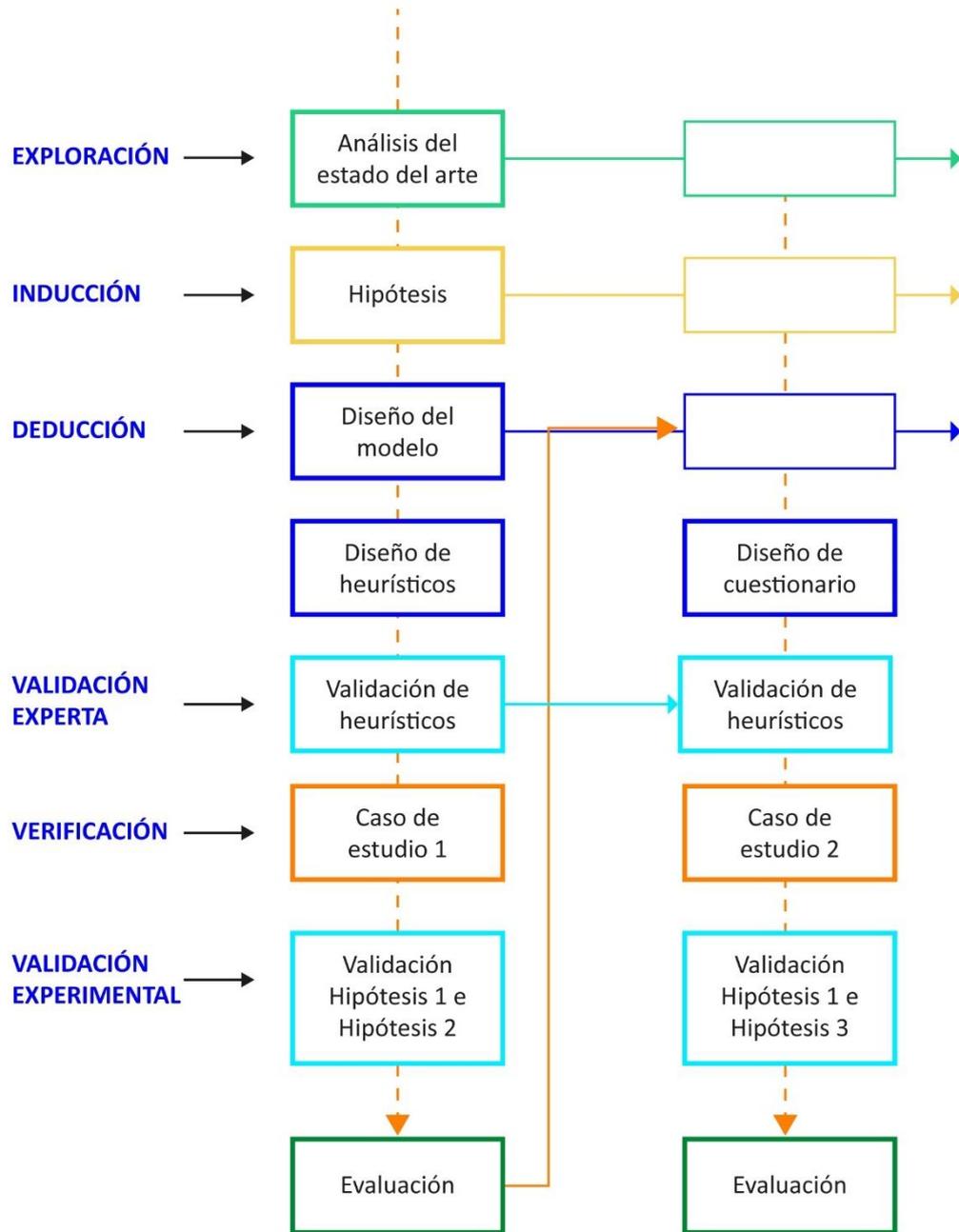


Figura 9: Metodología de investigación planteada siguiendo el modelo DIR (Horváth, 2007).

1.6 Estructura de la tesis

Esta tesis doctoral se compone de siete capítulos que abordan diferentes aspectos relacionados con la evaluación de la HRI en entornos industriales. A continuación, se proporciona una descripción general de cada uno de los capítulos.

Capítulo I: INTRODUCCIÓN

En el Capítulo I, se establece el objeto y el contexto de la tesis, se plantean las hipótesis de investigación, se definen los objetivos perseguidos y se describe la metodología de investigación utilizada.

Capítulo II: ENMARQUE CIENTÍFICO-TECNOLÓGICO

En el Capítulo II, se realiza un análisis exhaustivo del estado del arte en cuanto a la HRI en la nueva era industrial, incluyendo la revisión de modelos de evaluación en HRI, dispositivos fisiológicos utilizados para medir la UX (UX) y cuestionarios para la evaluación de la UX en entornos industriales. Además, se lleva a cabo un estudio crítico del estado del arte.

Capítulo III: MODELO ITPX-ROBOTS

En el Capítulo III se presentan los fundamentos teóricos y el marco conceptual de la tesis, así como el modelo propuesto para la evaluación de la HRI en entornos industriales. Se describe en detalle el modelo HRIX, el modelo HRI FIT y la taxonomía HUBOXT, que incluye características individuales, características del robot y características del contexto.

Capítulo IV: DESARROLLO Y VALIDACIÓN DE LA HERRAMIENTA DE EVALUACIÓN HEURÍSTICA HEUROBOX

En el Capítulo IV, se detalla el proceso de desarrollo y validación de la herramienta de evaluación heurística HEUROBOX. Se describen los requisitos para un nuevo listado de heurísticos, se presenta la metodología utilizada y se muestra el listado completo de heurísticos, organizados por categorías. También se presentan los resultados de la validación y priorización de los expertos, así como el indicador HEUROBOX.

Capítulo V: DESARROLLO Y VALIDACIÓN DEL CUESTIONARIO HUMAN-ROBOT EXPERIENCE (HUROX)

En el Capítulo V, se describe el proceso de desarrollo y validación del cuestionario Human-Robot Experience (HUROX). Se detalla la metodología utilizada, se presentan los ítems iniciales y se muestran los resultados de la validación aparente, el ratio de validez del contenido y el índice de validez del contenido. Finalmente, se presenta el cuestionario final.

Capítulo VI: VALIDACIÓN DEL MODELO ITPX-ROBOTS

En el Capítulo VI se lleva a cabo la validación del modelo propuesto, denominado ITPX-ROBOTS, a través de dos casos de estudio. Se describe la metodología utilizada, se presentan los resultados de cada caso de estudio y se realiza un estudio crítico de los resultados obtenidos.

Capítulo VII: CONCLUSIONES Y LÍNEAS FUTURAS

Finalmente, en el Capítulo VII, se presentan las conclusiones generales de la tesis, se evalúa la validación de las hipótesis planteadas, se destacan las aportaciones más relevantes y se discuten las limitaciones del estudio. Además, se sugieren posibles líneas de investigación futuras relacionadas con la evaluación de la HRI en entornos industriales.

Capítulo II

Enmarque científico- tecnológico

2 ENMARQUE CIENTÍFICO-TECNOLÓGICO

La presente investigación se centra en la HRI con un enfoque basado en el Diseño Centrado en las Personas (HCD, por sus siglas en inglés) y la UX. En este capítulo, se identifican los factores que influyen en la HRI en los entornos industriales y se exploran las herramientas y métricas para evaluar la percepción de las personas. La recogida de medidas objetivas y subjetivas es importante para comprender la situación del participante. Las herramientas fisiológicas proporcionan una evaluación objetiva del estado del participante, mientras que los cuestionarios ofrecen una medida subjetiva de su percepción.

En este capítulo, se presenta la revisión del estado actual de la investigación en lo que respecta a la interacción entre personas y robots en el ámbito industrial. Este análisis se lleva a cabo mediante una Revisión Sistemática de la Literatura (SLR, por sus siglas en inglés). A partir de esta revisión, se profundiza en el estudio de las diversas medidas y métricas que posibilitan la evaluación de la UX y la aceptación tecnológica en el contexto de la HRI.

A partir de esta revisión, se lleva a cabo un análisis crítico del estado del arte, siendo la base fundamental para el diseño del nuevo modelo propuesto. Este análisis crítico permite la identificación de las fortalezas y limitaciones de las investigaciones previas, así como de las áreas no exploradas en el campo de la HRI en entornos industriales. Esta revisión exhaustiva del estado del arte proporciona una sólida base teórica y contextual para el desarrollo del nuevo modelo de evaluación.

2.1 HRI en la nueva industria

Con la evolución de los entornos industriales hacia la implementación de fábricas inteligentes y la adopción de la fabricación ágil, así como el avance del IoT en los métodos de fabricación, se ha observado una creciente interconexión e interdependencia de herramientas, sistemas y procesos (Argyle et al., 2021). En este contexto, la integración de robots en los equipos humanos plantea desafíos únicos, tales como la falta de experiencia o conocimientos robóticos por parte de las personas, problemas de facilidad de uso o programación, y una generalizada desconfianza hacia los robots y sus capacidades (Argyle et al., 2021). Para hacer frente a estos retos, es fundamental que la interacción con un robot sea segura, natural y eficaz, al nivel de la interacción entre personas.

Para optimizar la HRI se requiere una comprensión más clara de las construcciones cognitivas sociales (Warta et al., 2016). La percepción de los robots ha evolucionado de considerarlos como herramientas que amplían las capacidades humanas a concebirlos como compañeros y colaboradores en equipos (Morrow & Fiore, 2012; Schaefer et al., 2017; T. Wiltshire, 2013; T. J. Wiltshire et al., 2017). Según Warta et al. (2016), para que los robots operen de esta manera, deben adquirir procesos cognitivos sociales complejos.

En la actualidad, sin embargo, en los equipos de colaboración persona-robot, las funciones de la persona operaria suelen limitarse a la supervisión, programación y mantenimiento. La supervisión implica la observación del estado de los equipos y la intervención periódica para la inspección, la asignación de tareas o el mantenimiento del proceso. Se trata de un nivel perpetuo de la interacción. En este sentido, el rol de la persona operaria es el que suele asociarse con HRI, sin embargo, los roles de supervisión, programación y mantenimiento se ven más afectados por la HMI (Sebastián García & del Hoyo Delgado, 2002).

La robótica colaborativa aporta a los procesos de producción las ventajas derivadas de las propiedades características de los robots. Estas propiedades incluyen un alto nivel de

precisión, velocidad y repetibilidad. En combinación con la flexibilidad y las habilidades cognitivas de las personas, el proceso de HRI consigue una colaboración eficiente entre ambos (Galin & Meshcheryakov, 2019; Heyer, 2010).

Welfare et al. (2019) indican que para minimizar los posibles impactos negativos de la integración de equipos persona-robot, se han de mantener los factores humanos de la interacción social, la autonomía y la resolución de problemas para garantizar la aceptación de la persona operaria y su experiencia positiva. En términos generales, a pesar de los avances en autonomía, los robots tienen más habilidades para tomar decisiones en ciertas áreas que en otras. Por ejemplo, la planificación estructurada (para la que existen algoritmos de procesos bien definidos) ha demostrado ser apta para la automatización. Sin embargo, la toma de decisiones no estructuradas sigue siendo el dominio de las personas, especialmente cuando se requiere sentido común (Clarke, 1994).

Prati et al. (2021) destacan que un aspecto clave para el éxito de la HRI es centrarse en la UX. En este contexto, la investigación sobre los robots sociales ha reconocido recientemente el papel del diseño de la UX para garantizar una interacción positiva. En efecto, una experiencia negativa podría impedir aprovechar las características de los robots o rechazar la interacción con ellos. Según Prati et al. (2021) solo el estudio preciso de la UX desde la fase inicial del diseño puede garantizar una HRI aceptable y agradable, capaz de mejorar también la eficiencia y la eficacia durante la interacción. Considerar la forma en que el operario interactúa con el robot y entender su UX es difícil, especialmente en el sector industrial. Para ello, es conveniente adoptar un enfoque estructurado centrado en las personas para ayudar a los diseñadores a resolver las cuestiones técnicas teniendo en cuenta las necesidades y capacidades de la persona usuaria, utilizando técnicas basadas en la UX

2.1.1 Relaciones persona-robot

Durante los próximos años aumentará la coexistencia entre personas y robots (Karabegović et al., 2020, 2022; Marvel et al., 2020). Esta coexistencia se dará en entornos tecnológicamente enriquecidos, en los que la información se deberá intercambiar de forma “natural” entre personas y robots, dando lugar a entornos híbridos en los que las personas coexisten en el mundo digital y el mundo real (Sebastián García & del Hoyo Delgado, 2002). En este sentido, la combinación de habilidades humanas y robóticas está empezando a cobrar cada vez más importancia (X. V. Wang et al., 2017). En el ámbito de la automatización, mientras ciertas tareas rutinarias o habilidades específicas pueden ser eficazmente apoyadas por sistemas automatizados, las decisiones locales o intervenciones excepcionales a menudo requieren la intervención humana debido a las características únicas de la situación, la complejidad involucrada o la naturaleza implícita del conocimiento necesario para encontrar una solución viable en un tiempo limitado y con recursos limitados disponibles. Hasta ahora, la colaboración entre recursos humanos y artificiales no ha sido una práctica común en la automatización, donde los robots y las personas suelen mantenerse separados y las personas deben adherirse a procedimientos de trabajo rígidos establecidos por el entorno de producción automatizado.

Sin embargo, la colaboración simbiótica entre personas y robots va más allá de estas limitaciones y requiere un entorno más flexible, transparente y accesible (X. V. Wang et al., 2017). Esto implica una mayor interacción entre personas y máquinas, donde ambos trabajan juntos de manera cooperativa, complementándose mutuamente para lograr objetivos comunes y resolver problemas de manera más efectiva. Esta evolución hacia una mayor colaboración entre personas y robots podría desbloquear nuevas oportunidades y beneficios en diversos campos, permitiendo una mejor adaptación a situaciones cambiantes y una mayor eficiencia en la realización de tareas complejas y críticas. Para lograr una colaboración simbiótica exitosa, es fundamental desarrollar sistemas y entornos que fomenten una

interacción fluida y natural entre personas y robots, además de garantizar la seguridad y la confianza en la cooperación entre ambos actores.

2.1.1.1 Taxonomía en HRI

La relación entre personas y robots en un entorno de trabajo compartido es un fenómeno polifacético que se clasifica según distintos puntos de vista (Kaiser et al., 2018). Diferentes taxonomías han sido propuestas para clasificar las relaciones entre personas y robots, y en este contexto, Yanco & Drury (2004) han desarrollado una taxonomía que considera diversos criterios para esta clasificación. Estos criterios son los siguientes:

- i) El tipo de tarea.
- ii) La criticidad de la tarea, es decir, el grado de importancia o urgencia de su realización, que puede ser alto, medio o bajo.
- iii) La morfología del robot, que se puede clasificar como antropomórfica, zoomórfica o funcional, según su apariencia o forma.
- iv) La ratio de personas por robot, que se representa mediante una fracción no reducida, indicando el número de personas involucrados en relación con el número de robots presentes en el entorno.
- v) La composición del equipo de robots, distinguiendo entre equipos homogéneos, donde todos los robots son del mismo tipo, y equipos heterogéneos, donde hay diferentes tipos de robots trabajando conjuntamente.
- vi) El nivel de interacciones compartidas entre grupos de personas y robots, es decir, el grado de colaboración y comunicación entre ellos durante la realización de tareas.
- vii) El tipo de proximidad física entre personas y robots, refiriéndose a la distancia física y contacto que ocurre entre ambos actores durante el trabajo conjunto.
- viii) El apoyo a las decisiones de las personas operarias, considerando si los robots asisten o complementan el proceso de toma de decisiones de las personas.
- ix) La temporalidad de las interacciones, que puede ser sincrónica, cuando personas y robots trabajan en tiempo real, o asíncrona, cuando sus acciones ocurren en momentos diferentes.
- x) La disposición espacial, determinando si personas y robots se encuentran en el mismo lugar físico o en lugares diferentes durante su interacción.
- xi) El nivel de autonomía de los robots, referente al porcentaje de tiempo en el que un robot realiza tareas de forma independiente, y la cantidad de intervención humana requerida, representando el porcentaje de tiempo que una persona operaria debe supervisar o controlar al robot.

Schmidtler et al. (2015) proponen una clasificación que se fundamenta en cuatro criterios distintos para analizar las interacciones entre personas y robots:

- i) El **espacio de trabajo**, considerando si la persona y el robot comparten un espacio de trabajo común o no.
- ii) El **tiempo de trabajo**, se define como el tiempo que el participante está trabajando dentro del espacio de trabajo.
- iii) El **objetivo**. Cada entidad del equipo que interactúa tiene un objetivo que quiere lograr. Este objetivo, puede coincidir o no con el de los demás integrantes del equipo.
- iv) El **contacto**.

En esta línea, Schmidtler et al. (2015) propusieron que la HRI puede categorizarse en tres tipos distintos (Tabla 2):

- i) **Coexistencia persona-robot**, lo que denota dos entidades, persona y robot, operando simultáneamente en un espacio de trabajo compartido.
- ii) **Cooperación persona-robot**, lo que significa que ambas entidades trabajan juntas en el mismo entorno hacia un objetivo común.
- iii) **Colaboración persona-robot** (HRC, por sus siglas en inglés), que además requiere una interacción física directa entre las dos entidades.

Tabla 2: Clasificación de la HRI en términos de espacio de trabajo, tiempo de trabajo, objetivo y contacto. Adaptado de Schmidtler et al. (2015).

	Coexistencia	Cooperación	Colaboración
Espacio de trabajo	✓	✓	✓
Tiempo de trabajo	✓	✓	✓
Objetivo		✓	✓
Contacto			✓

Siguiendo la clasificación propuesta por Schmidtler et al. (2015), X. V. Wang et al. (2017) hacen hincapié en tres características fundamentales para la categorización de la interacción entre personas y robots:

- i) **La relación temporal y espacial** de las personas y los robots, considerando si comparten un mismo espacio de trabajo y si interactúan simultáneamente o en momentos distintos.
- ii) **La multiplicidad de agentes**, donde se distingue entre agentes individuales, múltiples y en equipo. Los agentes en equipo actúan en base a un consenso y mantienen interacciones tanto con el entorno como con otros agentes de una manera coordinada y específica (Figura 10).

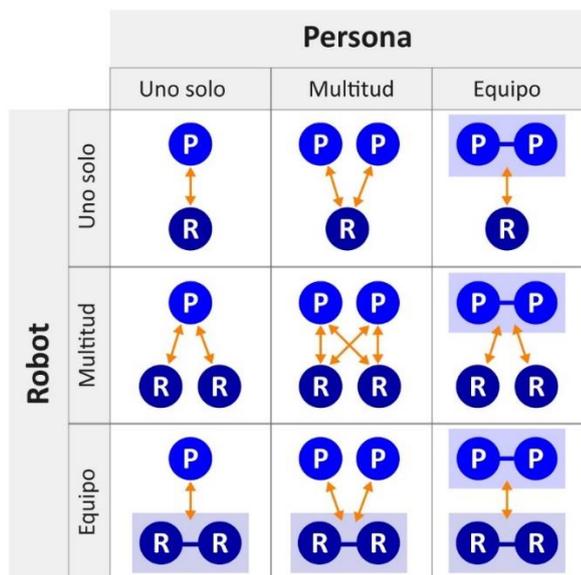


Figura 10: Esquema de clasificación de la colaboración entre personas y robots en función de la multiplicidad de los agentes. Adaptado de X.V. Wang et al. (2017).

- iii) **La autonomía de los agentes y las relaciones estrechamente vinculadas entre líder y seguidor** expresan la medida en que la acción del robot está directamente determinada por las personas o qué persona toma la iniciativa en una tarea específica (Bruemmer et al., 2005). Esta división en términos de

autonomía o iniciativa puede variar según el campo de aplicación. En un contexto industrial, se pueden distinguir comportamientos inactivos (descanso), activos (liderazgo) y de apoyo. Es importante mencionar que muchas de estas consideraciones asumen que estos roles se asignan previamente a la ejecución de la tarea (Figura 11).

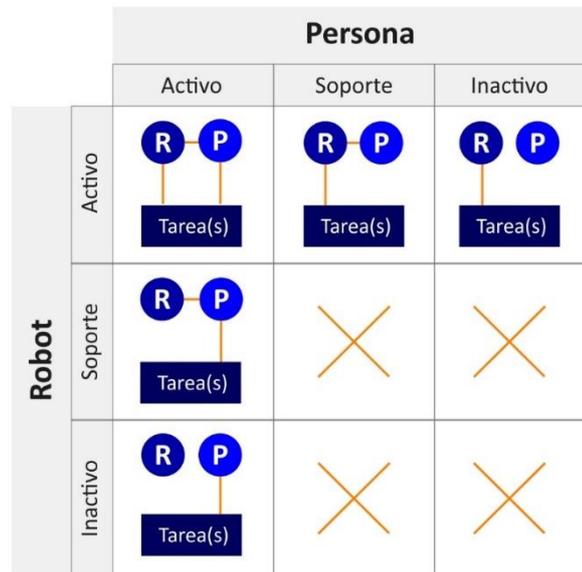


Figura 11: Esquema de clasificación de la colaboración entre personas y robots en función de la iniciativa. Adaptado de X.V. Wang et al. (2017).

A pesar de que toda la taxonomía es relevante desde el punto de vista del diseño, Sebastián García & del Hoyo Delgado (2002) y X. V. Wang et al. (2018) consideran la existencia de las siguientes dimensiones como las más relevantes para clasificar las relaciones entre personas y robots. Estas dimensiones son:

- i) El **espacio de trabajo** compartido se refiere en general, a si la persona y el robot trabajan en la misma área de trabajo sin vallas físicas o virtuales de separación.
- ii) El **contacto directo**, que indica si existe contacto físico directo entre el robot y la persona.
- iii) La **tarea de trabajo** compartida, que representa si la persona y el robot trabajan en la misma operación hacia el mismo objetivo de trabajo.
- iv) El **proceso simultáneo**, que significa que la persona y el robot trabajan al mismo tiempo, pudiendo ser en una misma tarea o en tareas diferentes.
- v) El **proceso secuencial**, que indica que las operaciones de la persona y el robot se disponen una tras otra sin solaparse en la escala temporal.

En consecuencia, la clasificación de las relaciones entre persona y robot pueden resumirse de la siguiente manera (Tabla 3) según L. Wang et al. (2019) y X. V. Wang et al. (2018):

- **Coexistencia:** Cuando un robot y una persona se sitúan dentro del espacio físico, pero sin solapar el espacio de trabajo del otro. No hay contacto directo entre la persona y el robot. El objeto de trabajo puede intercambiarse entre ellos, pero el proceso se realiza de forma independiente y simultánea.
- **Interacción:** Se produce cuando una persona y un robot que comparten el mismo espacio de trabajo se comunican entre sí. Una de las partes guía o controla a la otra, o se produce algún contacto físico (planificado o no) entre ellas. Tanto la persona como

el robot pueden trabajar en la misma tarea, pero la completan paso a paso en un orden secuencial.

- **Cooperación:** Puede desarrollarse entre personas y robots que tienen su propia autonomía. Con la esperanza de obtener un beneficio mutuo, los agentes que cooperan pueden compartir temporalmente algunos de sus recursos físicos, cognitivos o computacionales, aunque persigan sus propios intereses. Las partes pueden compartir un espacio de trabajo parcialmente superpuesto, pero no es típico el contacto directo entre ellas. Pueden trabajar simultáneamente, pero a veces tienen que esperar a la disponibilidad del otro agente o agentes.
- **Colaboración:** Es la actividad conjunta de personas y robots en un espacio de trabajo compartido, con el objetivo definido de realizar juntos un conjunto de tareas de trabajo determinadas. Suele requerir una actividad coordinada y sincrónica de todas las partes en la que también se permite el contacto físico. En cualquier caso, la colaboración supone una actividad conjunta y orientada a objetivos de las partes que comparten sus diferentes capacidades, competencias y recursos.

Tabla 3: Clasificación según las características de las relaciones persona-robot. Adaptado de L. Wang et al. (2019).

	Coexistencia	Interacción	Cooperación	Colaboración
Espacio de trabajo		✓	✓	✓
Contacto directo		✓		✓
Tarea		✓		✓
Fuente			✓	✓
Proceso simultáneo	✓		✓	✓
Proceso secuencial		✓	✓	

Además, X. V. Wang et al. (2018) proporcionan una síntesis sobre la relación entre la interacción y el nivel de riesgo, como se muestra en la Figura 12. En el caso de la coexistencia persona-robot, el compromiso y la complejidad de la interacción son bajos, y la seguridad es relativamente fácil de garantizar debido a la presencia de límites físicos que protegen al operador. En la HRI, el riesgo de seguridad aumenta considerablemente debido a que la persona y el robot comienzan a compartir el mismo espacio físico. Por otro lado, en el modo de cooperación persona-robot, tanto la interacción como el riesgo de seguridad aumentan considerablemente debido al mayor contacto directo entre la persona operaria y el robot. Sin embargo, la colaboración totalmente simbiótica entre la persona y el robot implica que ambas partes realizan la tarea de manera colaborativa, lo que conlleva inevitablemente un contacto directo entre la persona operaria y el robot. Por lo tanto, el nivel de riesgo de seguridad es mayor en comparación con el modo de cooperación.

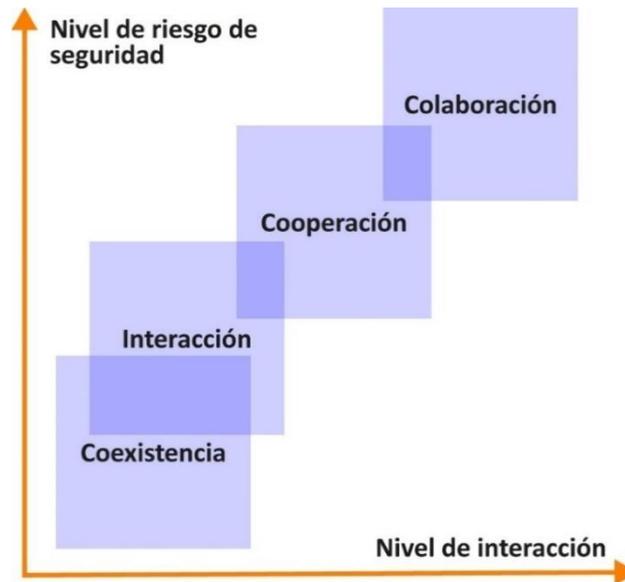


Figura 12: Interacción y nivel de riesgo de los distintos tipos de relación entre personas y robots. Adaptado de X.V. Wang et al. (2018).

Conforme a las afirmaciones de X. V. Wang et al. (2017), desarrollar soluciones para la colaboración persona-robot implica un marco de análisis y síntesis que contemple: (i) medios para clasificar y caracterizar el problema, y (ii) plantillas de soluciones y directrices para elaborar una solución que se ajuste adecuadamente a las premisas de producción existentes. En este contexto, los elementos fundamentales en un escenario de HRC son los siguientes:

- i) **Los agentes**, es decir, los robots y las personas que participan activamente en el proceso de producción o que ocupan una parte del espacio disponible, junto con las multiplicidades y los roles que perfilan dónde es necesario establecer prioridades o resolver conflictos. Las características de los agentes también incluyen los modos de comunicación y detección, la geometría, las propiedades cinemáticas y dinámicas, las características de temporización y, especialmente en el caso de las personas, las capas de acciones posibles.
- ii) El **entorno de trabajo** incluye los recursos necesarios para la producción pero que no desempeñan un papel activo y las condiciones (luz ambiental, ruido, etc.) que afectan a las soluciones viables.
- iii) **Las piezas y las operaciones** suelen estar en interdependencia inseparable. Ambas pueden tener variaciones que deben manejarse durante la producción. Incluso un conjunto inequívoco de componentes puede tener múltiples órdenes de montaje, y la elección concreta puede depender de los agentes y los recursos disponibles.

Los robots, como la mayoría de los artefactos tecnológicos, requieren interfaces de usuario para interactuar con las personas. Las interfaces desempeñan un papel central, siendo el canal principal de comunicación. Una interacción entre personas y robots de alta calidad requiere interfaces intuitivas para la persona usuaria. Por un lado, la persona operaria puede dar a los robots entradas de manera sencilla, sin ninguna distracción de su tarea principal; por otro lado, el robot proporciona información más clara al usuario, generando una comprensión e interpretación de datos inmediata (Sebastián García & del Hoyo Delgado, 2002). La adopción de interfaces intuitivas se hace aún más importante en el caso de una colaboración más estrecha entre los robots y las personas.

Según la tipología de la comunicación, las interfaces persona-robot pueden clasificarse en cuatro categorías (Goodrich & Schultz, 2008): (i) Pantallas visuales (por ejemplo, interfaces

gráficas de usuario, interfaces de realidad aumentada), (ii) gestuales (por ejemplo, movimientos de la mano y la cara), (iii) voz y lenguaje natural (por ejemplo, respuestas auditivas y basadas en texto), y (iv) interacciones físicas y hápticas.

Según Prati et al. (2021), estas interfaces también pueden estar relacionadas con el nivel de interacción proporcionado. En concreto, el primer nivel de interacción (coexistencia) suele satisfacerse con interfaces gráficas. El segundo nivel de interacción (cooperación) suele requerir interfaces más avanzadas, como las de voz y gestos. Por último, el tercer nivel de interacción (colaboración) puede requerir también una interacción física o háptica directa para que sea eficaz y natural al mismo tiempo. La Figura 13 muestra el marco propuesto por Prati et al. (2021) para clasificar la HRI, teniendo en cuenta los tres niveles de interacción y las relaciones con los distintos tipos de interfaz.

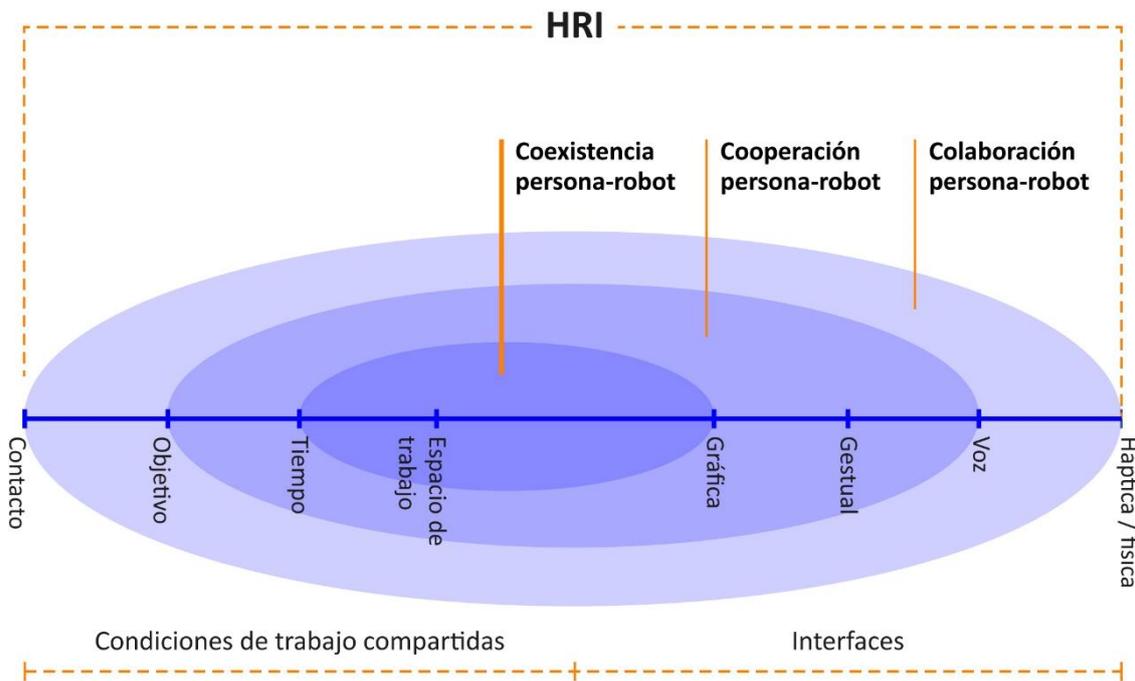


Figura 13: Marco para clasificar la HRI: niveles de interacción y relaciones con los distintos tipos de interfaz. Adaptado de Prati et al. (2021).

2.2 Modelos de evaluación en HRI: Revisión Sistemática de la Literatura (SLR)

Numerosas contribuciones se han escrito sobre el diseño y la evaluación de la HRI. La adopción reciente del concepto "Industria 5.0" por parte de la Comisión Europea (2021) ha aumentado el interés en incorporar factores humanos. Sin embargo, la literatura reporta pocos intentos de integrar métricas de factores humanos de manera integral para evaluar la UX en HRI.

Existen varias revisiones de literatura en el contexto de la HRI. El trabajo de Hentout et al. (2019) propone una clasificación general del contenido de trabajos en HRI en varias categorías y subcategorías, como el diseño de *hardware* y *software* de sistemas robóticos colaborativos, seguridad en la robótica industrial e HRI cognitiva. Señalaron que el objetivo de la HRI es proporcionar a los robots tres requisitos fundamentales: (i) la intención humana debe ser fácil de inferir por el robot, (ii) el control debe ser intuitivo desde el punto de vista de la persona y (iii) el controlador diseñado debe ser seguro tanto para personas como para robots. Simões et al. (2022) enumeraron una serie de pautas clasificadas en general en: (i) persona operaria y tecnología, (ii) rendimiento del equipo persona-robot y (iii) enfoque integrado para

el diseño de HRI. Como conclusión general, destacaron la importancia del *feedback* para mejorar la confianza y la atribución de culpa. Presentaron recomendaciones para el diseño de lugares de trabajo centrados en las personas, seguros, ergonómicos y saludables, donde se consideran aspectos técnicos, sociales y psicofísicos de la colaboración. Savela et al. (2018) analizaron cómo se ha estudiado la aceptación social de los robots en diferentes campos ocupacionales y qué tipos de actitudes se han descubierto con respecto a los robots como trabajadores. Sus resultados implican que las actitudes hacia los robots son positivas en muchos campos de trabajo. Sin embargo, indicaron que se necesitan medidas validadas. Veling & McGinn (2021) analizaron el uso de métodos y enfoques cualitativos en la literatura de HRI para contribuir al desarrollo de una base de enfoques y metodologías en el área de investigación. Su revisión reveló seis métodos predominantes de recolección de datos cualitativos en la literatura de HRI: observaciones cualitativas, entrevistas semi-estructuradas, grupos de enfoque, actividades generativas, cuentas reflexivas y narrativas y análisis de contenido textual.

Se han publicado numerosos estudios sobre robots en la industria, pero la mayoría de ellos se centran en aspectos de seguridad y protección (Gopinath et al., 2017; Hentout et al., 2019; Tsarouchi et al., 2016). Otros investigadores han estudiado la estandarización para mejorar los lugares de trabajo (Gualtieri, Monizza, et al., 2020; Tsai et al., 2014). En este sentido, la industria de la robótica está creciendo hasta un nivel en el que las personas y los robots podrán colaborar (Harriott et al., 2013). Sin embargo, como señalaron Harriot et al. (2013), todavía no existe un modelo universal que evalúe el efecto de esta colaboración en el rendimiento de las personas.

Además, se observa que en la literatura no se ha prestado atención a los factores humanos resultantes de la HRI. Los factores emocionales, como la confianza, la satisfacción o la carga mental, han sido poco estudiados para la optimización de sistemas robóticos colaborativos. La evaluación de estos factores es beneficiosa para conocer cómo se sienten las personas antes, durante y después de la interacción. De esta manera, las acciones del robot podrían adaptarse a las necesidades de las personas, en línea con el enfoque HCD.

La dirección indicada por el marco Horizon 2020 de la Comisión Europea establece que la igualdad de género debe ser promovida a través de cambios en la cultura de las instituciones científicas, cambios en la composición de los equipos de investigación para lograr la igualdad y cambios en el contenido y diseño de las actividades de investigación (Díaz, 2021). Sin embargo, la perspectiva de género no se ha incluido en ninguno de los estudios identificados. La integración de la perspectiva de género en la investigación es necesaria para evitar sesgos, donde a veces se construyen como norma las realidades, experiencias y expectativas de un grupo de personas (considerando a los hombres como referencia), produciendo así resultados parciales y no universales. La perspectiva de género en la investigación implica integrar variables de sexo y género en el proceso científico, lo que tendrá implicaciones al considerar las normas de género, identidades y relaciones como variables explicativas del fenómeno analizado. Por otro lado, muchas personas enfrentan barreras de accesibilidad al interactuar con robots, principalmente personas que no suelen interactuar con nuevas tecnologías, personas mayores y personas usuarias con discapacidades (Qbilat & Iglesias, 2018). Diseñar y desarrollar sistemas robóticos que garanticen la accesibilidad a todas las personas usuarias con diferentes habilidades y necesidades es esencial para hacer que los sistemas de HRI sean más inclusivos. Sin embargo, ninguna de las revisiones identificadas considera la inclusividad como un aspecto necesario para abordar.

Objetivo de la Revisión Sistemática de la Literatura

El objetivo de esta Revisión Sistemática de la Literatura (SLR) es identificar evaluaciones de la HRI que incluyan una perspectiva HCD y comprender los factores humanos que afectan en la interacción entre la persona y el robot en un entorno industrial.

2.2.1 Metodología

En este método de revisión de la literatura, se utilizan las directrices propuestas por Kitchenham & Charters (2007) para llevar a cabo la SLR en este trabajo, que incluye cinco fases. Estas fases tienen sus propias subdivisiones: (1) la planificación, el alcance y la definición de las preguntas de investigación, (2) la búsqueda de la literatura en base a la planificación, es decir, la definición de fuentes de datos/información y la búsqueda en ellas, (3) la selección de la literatura en base a criterios de inclusión/exclusión y al estudio de evaluación de la calidad, (4) el resumen, extracción y síntesis de los resultados y, (5) la distribución de los resultados. Las siguientes secciones describen estas fases de la metodología de revisión.

2.2.1.1 Planificación y alcance

En esta fase se establecen los fundamentos y criterios necesarios para abordar la SLR. A continuación, se presentan las preguntas de investigación que guiarán el proceso de búsqueda y análisis de la literatura relevante, así como las fuentes de información utilizadas para llevar a cabo dicha revisión.

2.2.1.1.1 Preguntas de investigación

En esta revisión, se han formulado dos Preguntas de Investigación (PI) que se muestran en la Tabla 4.

Tabla 4: Preguntas de investigación para la SLR.

Id	Pregunta de investigación
PI-1	¿Existe un modelo de evaluación de la HRI que incluya los factores humanos en entornos industriales?
PI-2	¿Qué factores humanos incluye y cómo los evalúa?

2.2.1.1.2 Fuentes de información

Para la búsqueda de información se han empleado los siguientes términos clave:

1. "Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Workstation"
2. "Human factors" OR "user experience" OR "UX"
3. Evaluation OR Assessment

De este modo, se ha creado la ecuación de búsqueda principal:

("Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR "UX") AND (Evaluation OR Assessment)

Para la búsqueda de trabajos, se han utilizado diversas bases de datos electrónicas, que se detallan en la Tabla 5.

Tabla 5: Bases de datos utilizadas en la SLR.

Nombre de la Base de Datos	Tipo de Base de Datos	Descripción
ACM Digital Library	Base de datos de investigación	Base de datos científica en materias relacionadas con la informática y ciencias computacionales.
Engineering Village	Base de datos de investigación	Base de datos especializada en ingeniería.
IEEE Xplore Digital Library	Base de datos de investigación	Base de datos científica en materias relacionadas con la informática, la ingeniería eléctrica y la electrónica.
Inspec	Base de datos especializada	Base de datos especializada en: física, ingeniería eléctrica y electrónica, computación y control, tecnología de la información, e ingeniería mecánica y de producción.
Science Direct	Base de datos de la editorial	Ofrece sus propias revistas científicas a texto completo. Cubre la mayoría de las disciplinas, pero con un enfoque en la ciencia, la tecnología y ciencias sociales, de Elsevier y otras editoriales relacionadas.
Scopus	Base de datos de citas	Servicio de indexación de citas científicas para la búsqueda de citas de artículos de revistas revisadas por expertos. Se utiliza principalmente en los cálculos bibliométricos, de Elsevier.
Web of Science	Base de datos de citas	Servicio de indexación de citas científicas para la búsqueda de citas de artículos de revistas revisadas por expertos. Se utiliza principalmente en los cálculos bibliométricos, de Thomson/Reuter.

2.2.1.2 Búsqueda de la literatura

En el marco de esta investigación, se ha procedido a realizar la búsqueda de la literatura en diversas bases de datos, ajustando la ecuación de búsqueda según las características y requisitos de cada base de datos en particular (Tabla 6). Se han establecido ciertos criterios generales para la selección de artículos, incluyendo que estos sean artículos de revistas evaluadas por expertos y que estén publicados en el período comprendido entre enero de 2011 y la fecha de la búsqueda (julio de 2021). La elección de revistas evaluadas por pares asegura la validez y calidad del conocimiento presentado en los artículos (Kerr & Coviello, 2019), mientras que el límite de años de publicación se estableció para concentrar la búsqueda en trabajos más recientes y relevantes. Se presume que la literatura anterior a 2011 es menos relevante para los objetivos de esta revisión.

Tabla 6: Ecuación de búsqueda utilizada en cada una de las bases de datos.

N.º	Base de datos	Fecha en la que se ha realizado la búsqueda	Ecuación de búsqueda	N.º de resultados
1	ACM	05/07/2021	[[All: "human-robot interaction"] OR [All: "human-robot collaboration"] OR [All: "human-robot coexistence"] OR [All: "human-robot cooperation"] OR [All: "human-robot workstation"]] AND [[All: "human factors"] OR [All: "user experience"] OR [All: ux]] AND [[All: "evaluation"] OR [All: "assessment"]] AND [Publication Date: (01/01/2011 TO 12/31/2021)]nt")	67
2	Engineering Village	12/07/2021	(((((("Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Cooperation" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR ux) AND ("Evaluation" OR "Assessment")) WN ALL)) AND (({ja} WN DT) AND ((2021 OR 2020 OR 2019 OR 2018 OR 2017 OR 2016 OR 2015 OR 2014 OR 2013 OR 2012 OR 2011) WN YR)))	29
3	IEEE Xplore	12/07/2021	("Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Cooperation" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR ux) AND ("Evaluation" OR "Assessment")	20
4	Inspec		("Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Cooperation" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR ux) AND ("Evaluation" OR "Assessment")	53
5	ScienceDirect		5.1 ("Human-Robot Interaction" OR "Human-Robot Collaboration") AND ("HUMAN FACTORS" OR "USER EXPERIENCE") AND (EVALUATION or assessment) 5.2 ("Human-Robot Coexistence" OR "Human-Robot Cooperation" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR ux) AND ("Evaluation" OR "Assessment")	296
6	Scopus		TITLE-ABS-KEY (("Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Cooperation" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR ux) AND ("Evaluation" OR "Assessment")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015) OR LIMIT-TO (PUBYEAR , 2014) OR LIMIT-TO (PUBYEAR , 2013) OR LIMIT-TO (PUBYEAR , 2012) OR LIMIT-TO (PUBYEAR , 2011)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))	51
7	Web of Science		("Human-Robot Interaction" OR "Human-Robot Collaboration" OR "Human-Robot Coexistence" OR "Human-Robot Cooperation" OR "Human-Robot Workstation") AND ("Human factors" OR "user experience" OR ux) AND ("Evaluation" OR "Assessment")	39

2.2.1.3 Selección de la literatura

En esta fase, se detallan los criterios de inclusión y exclusión establecidos. Asimismo, se explican las preguntas de evaluación de calidad que se han utilizado para evaluar la pertinencia y utilidad de los estudios seleccionados.

2.2.1.3.1 Criterios de inclusión y exclusión

El proceso de selección de la literatura ha comenzado con la identificación y exclusión de artículos duplicados, lo cual ha resultado en un total de 117 artículos duplicados identificados y descartados.

Posteriormente, se ha llevado a cabo una selección más exhaustiva de la literatura, basada en el análisis del título y el resumen de los artículos. Para guiar esta selección, se han definido criterios específicos de inclusión y exclusión, los cuales se detallan en la Tabla 7. Durante este paso, se han excluido un total de 331 artículos que no cumplían con los criterios establecidos y que no se consideraban pertinentes para la revisión sistemática en cuestión.

Tabla 7: Criterios de inclusión y exclusión utilizados para la selección de la literatura.

Criterio	Inclusión	Exclusión
Idioma	Inglés o español	Otro que no sea inglés o español.
Fecha de publicación	De 2011 a 2021	Artículos que no se encuentran en el rango entre 2011 y 2021.
Preguntas de investigación	Artículos que pueden responder al menos a una de las PI	Artículos que no responden a ninguna PI.

2.2.1.3.2 Estudio de Evaluación de Calidad (EC)

El estudio de evaluación de calidad desplegado en esta revisión tiene como objetivo determinar la pertinencia y utilidad de los artículos seleccionados con respecto al tema específico abordado. Este proceso permite identificar si los artículos están relacionados con el tema específico que se está revisando y si son útiles al considerar una evaluación de la UX en un entorno robótico industrial.

Con este propósito, se formularon cinco preguntas de evaluación de calidad (EC), las cuales fueron revisadas en detalle y calificadas en función del análisis realizado (Tabla 8). Dado el carácter experimental de este tipo de investigación, el tamaño de muestra o replicabilidad no se ha establecido como un aspecto crucial en la EC. Existe una gran variabilidad en el tamaño de muestra utilizado en los diferentes experimentos debido a los distintos protocolos de aplicación que pueden existir. Dado que la EC no existe de manera aislada, sino que sirve directa o indirectamente para responder a las preguntas de investigación y respaldar las conclusiones de este estudio (L. Yang et al., 2021), estas preguntas de investigación están asociadas únicamente con los factores HRI, las formas de evaluarla y la consideración de género e inclusión. Por lo tanto, definimos las preguntas de EC siguiendo las recomendaciones de (L. Yang et al., 2021).

Tabla 8: Preguntas para la evaluación de la calidad.

EC	Preguntas de evaluación de calidad	Objetivo
EC1	¿El tema propuesto está relacionado con los factores humanos en una HRI en un entorno industrial?	Asignar puntajes más altos a los artículos relacionados con los factores humanos en la HRI, especialmente en entornos industriales.
EC2	¿Ayuda esta investigación a identificar los factores humanos que afectan a la HRI?	Valorar aquellos documentos que, al menos, contribuyen a identificar los factores humanos que influyen en la HRI, aquellos más alineados con el enfoque HCD, y, por ende, con el propósito de esta investigación.
EC3	¿Se describe adecuadamente el tema propuesto?	Evaluar artículos de calidad y rigor, que estén bien redactados y que describan adecuadamente el tema propuesto.
EC4	¿Si el tema propuesto tiene en cuenta una perspectiva de género o inclusiva?	Asignar mayor puntuación a aquellos artículos que consideran la perspectiva de género y la inclusión.
EC5	¿Describe esta investigación cómo evaluar la experiencia de la persona usuaria en la HRI?	Dar mayor puntuación aquellos artículos que presentan experimentos en los que se llevan a cabo evaluaciones de HRI desde la perspectiva de la experiencia de la persona usuaria.

Las cinco preguntas de evaluación de calidad mencionadas en la Tabla 8 juegan un papel fundamental en el proceso de revisión de los estudios seleccionados, ya que permiten analizar su contribución al presente estudio. El propósito de la evaluación de la calidad es facilitar la comprensión de la idoneidad y utilidad de estos estudios para nuestra investigación. En este contexto, Nidhra et al. (2013) propusieron criterios de calidad de alto nivel, que se traducen en puntuaciones específicas para cada hallazgo, asignándoles tres tipos de calificaciones: alta, media y baja. Estas calificaciones se otorgan en respuesta a las preguntas de evaluación de calidad planteadas.

La puntuación de 2 se asigna a aquellos estudios que cumplen completamente con los estándares de calidad establecidos. Los estudios que cumplen parcialmente con estos estándares obtienen una puntuación de 1, mientras que aquellos que no cumplen con los criterios de calidad reciben una puntuación de 0. Por lo tanto, la puntuación máxima posible para cada estudio es de 10 (es decir, $5 \times 2 = 10$). En caso de que un estudio no cumpla con los estándares de calidad, su puntuación será 0, representando la calificación más baja posible (es decir, $5 \times 0 = 0$).

En el contexto de esta SLR, se han tenido en cuenta aquellos artículos que obtuvieron una puntuación mayor que 7 (Tabla 9), lo que representa un resultado razonable y garantiza la calidad y confiabilidad de los hallazgos. En total, se seleccionaron 24 artículos que cumplían con este criterio.

Tabla 9: Puntuaciones de los artículos seleccionados.

Puntuación obtenida	N. de artículos	Referencia
9	8	Lindblom & Alenljung (2020); Daniel et al. (2013); Danielsson et al. (2017); Lasota & Shah (2015); Schillaci et al. (2013); von der Pütten & Bock (2018); Kildal et al. (2019); Meissner et al. (2020).
8	16	Colim et al. (2021); Prati et al. (2021); Harriott et al. (2013); Charalambous et al. (2017); Aromaa et al. (2018); Villani et al. (2018); Cohen et al. (2018); Tang & Webb (2018); Baskaran et al. (2019); Almeida et al. (2020); Hietanen et al. (2020); Gualtieri, Rauch, et al. (2020); Pantano et al. (2020); Beschi et al. (2020); Joosse et al. (2021); Qbilat et al. (2021).

2.2.1.4 Proceso de la revisión sistemática

En la Figura 14 se presenta el proceso llevado a cabo durante la revisión de la literatura. En la primera fase, se identificó un total de 555 artículos relevantes para el tema de estudio. Posteriormente, se detectaron 117 artículos duplicados, los cuales fueron excluidos de la revisión. En la segunda fase, se realizó un proceso de cribado mediante la lectura minuciosa de los títulos y *abstracts* de los 438 artículos restantes, aplicando los criterios de inclusión y exclusión previamente definidos. Como resultado, se excluyeron 331 artículos que no cumplían con los criterios establecidos.

En la siguiente etapa, se procedió a una lectura completa de los 107 artículos que superaron el proceso de cribado. Cada uno de estos artículos fue evaluado individualmente, utilizando las preguntas de evaluación de calidad previamente formuladas. Siguiendo este proceso de evaluación, se consideraron únicamente los artículos que obtuvieron una puntuación igual o mayor que 7 sobre 10. Como resultado, un total de 24 artículos cumplían con este criterio y fueron seleccionados para un análisis más detallado, en el cual se extrajeron datos relevantes y se llevó a cabo la síntesis de la información.

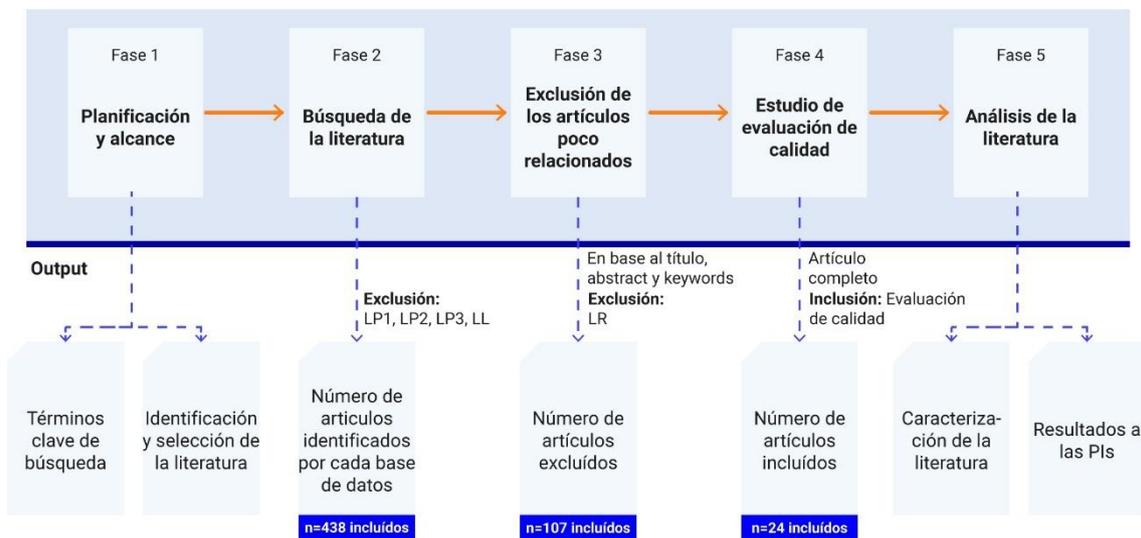


Figura 14: Proceso seguido en la Revisión Sistemática de la Literatura (SLR).

2.2.2 Resultados

A continuación, se presentan los resultados obtenidos tras el riguroso proceso de revisión de la literatura sobre la HRI en entornos industriales. Se abordan aspectos relevantes relacionados con los factores humanos, las metodologías y técnicas de evaluación, así como la perspectiva de género y la inclusividad en este contexto.

2.2.2.1 Caracterización de la literatura

La caracterización de permite comprender el estado actual del campo estudiado e identificar tendencias y cambios significativos a lo largo del tiempo.

2.2.2.1.1 Evolución en el campo

El análisis de la evolución en el campo de la HRI en entornos industriales revela una tendencia creciente en la producción de literatura científica en esta área. El primer artículo identificado data del año 2013, que fue seguido por un número limitado de publicaciones durante los años siguientes. En concreto, se identificaron tres artículos correspondientes a ese año: los estudios de Harriott et al. (2013), Schillaci et al. (2013) y Daniel et al. (2013). No es hasta 2017 que las publicaciones que estudian la evaluación de los factores humanos en

entornos persona-robot industriales comienzan a incrementarse. De hecho, más del 80% de las publicaciones identificadas datan del 2017 al 2021 (Tabla 9).

Sin embargo, a partir del año 2017, se observa un notable incremento en el número de publicaciones que abordan la evaluación de los factores humanos en entornos persona-robot industriales. De hecho, más del 80% de las publicaciones identificadas en esta revisión datan del periodo comprendido entre 2017 y 2021 (Tabla 10).

Este incremento sostenido en la producción de literatura durante los últimos años indica un creciente interés y relevancia de esta temática en la comunidad científica, lo cual sugiere una mayor comprensión de la importancia de la interacción entre personas y robots en el contexto industrial y las implicaciones que esto tiene en la mejora de la eficiencia, seguridad y experiencia de la persona usuaria.

Tabla 10: Número de artículos identificados por año.

Año	N.º	% distribución
2012	0	0,00%
2013	3	12,50%
2014	0	0,00%
2015	1	4,17%
2016	0	0,00%
2017	2	8,33%
2018	5	20,83%
2019	2	8,33%
2020	7	29,17%
2021	4	16,67%

2.2.2.1.2 Naturaleza de las revistas

Las revistas que han presentado una mayor cantidad de artículos identificados en esta revisión son *Procedia CIRP* y *ACM Transactions on Human-Robot Interaction (THRI)*, con tres artículos en cada una de ellas. *Procedia CIRP* es una revista especializada en la publicación de actas de alta calidad de las conferencias CIRP, mientras que *ACM Transactions on Human-Robot Interaction (THRI)* se centra en la investigación en el campo de la HRI. Además, se han identificado dos artículos en las revistas *IFAC PapersOnLine*, *Procedia Manufacturing* y *Robotics and Computer Integrated Manufacturing* (Tabla 11).

Tabla 11: Revistas en las que más artículos se han identificado.

Nombre de la revista	N.º de artículos identificados
Procedia CIRP	3
ACM Transactions on Human-Robot Interaction (THRI)	3
IFAC-PapersOnLine	2
Procedia Manufacturing	2
Robotics and Computer-Integrated Manufacturing	2

En cuanto al impacto de las publicaciones, se observa que el 96% de los artículos identificados pertenecen a revistas indexadas. De las 17 revistas identificadas, 5 de ellas (29%) se clasifican en el primer cuartil, mientras que 9 (53%) se encuentran en el segundo cuartil en términos de impacto (Tabla 12).

Tabla 12: Impacto de las revistas.

Revista	Cuartil	SJR 2020
ACM Transactions on Human-Robot Interaction (THRI)	Q2	0,6
Applied Sciences	Q2	0,44
Electronics	Q2	0,36
Human Factors	Q1	0,82
IFAC-PapersOnLine	Q3	0,31
International Journal of Advanced Manufacturing Technology	Q1	0,95
International Journal of Human-Computer Studies	Q2	0,73
International Journal of Social Robotics	Q1	0,66
Journal of Robotics	Q2	0,3
Mechatronics	Q1	0,94
Modeling, Identification and Control	Q4	0,22
Procedia CIRP	nd	0,68
Procedia Manufacturing	Q2	0,5
Proceedings of the ACM on Human-Computer Interaction	Q2	0,38
Robotics and Computer-Integrated Manufacturing	Q1	1,56
Sensors	Q2	0,64
Sustainability	Q2	0,61

2.2.2.1.3 Número de citas por artículo

En relación con el número de citas en los artículos revisados, se observa una cantidad relativamente baja en comparación con campos de investigación más amplios. El artículo más citado es el de Villani et al. (2018) con un total de 591 citas. En segunda posición se encuentra el trabajo de Lasota & Shah (2015) con 143 citas. El tercer artículo más citado es el de los autores Hietanen et al. (2020) con un total de 111 citas. La Tabla 13 presenta una lista con los diez artículos más citados incluidos en esta revisión.

Tabla 13: Artículos más citados.

Título	Autor/es	Año	Nº de citas (agosto 2023)
Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications	Villani, V., Pini, F., Leali, F., Secchi, C.	2018	591
Analyzing the Effects of Human-Aware Motion Planning on Close-Proximity Human–Robot Collaboration	Lasota, P., Shah, J.	2015	143
AR-based interaction for human-robot collaborative manufacturing	Hietanen, A., Pieters, R., Lanz, M., Latokartano, J., Kämäräinen, J.	2020	111
Digital Human and Robot Simulation in Automotive Assembly using Siemens Process Simulate: A Feasibility Study	Baskaran, S., Niaki, F., Tomaszewski, M., Gill, J., Chen, Y., Jia, Y., Mears, L., Krovi, V.	2019	37
Evaluating the Effect of Saliency Detection and Attention Manipulation in Human-Robot Interaction	Schillaci, G., Bodirosa, S., Hafner, V.	2013	31
The development of a Human Factors Readiness Level tool for implementing industrial human-robot	Charalambous, G., Fletcher, S., Webb, P.	2017	30

collaboration			
Assessing Instructions in Augmented Reality for Human-robot Collaborative Assembly by Using Demonstrators	Danielsson, O., Syberfeldt, A., Brewster, R., Wang, L.	2017	29
The Design and Evaluation of an Ergonomic Contactless Gesture Control System for Industrial Robots	Tang, G., Webb, P.	2018	24
Empowering assembly workers with cognitive disabilities by working with collaborative robots: a study to capture design requirements	Kildal, J., Martín, M., Ipiña, I., Maurtua, I.	2019	24
Assessing physical workload for human-robot peer-based teams	Harriott, C., Zhang, T., Adams, J.	2013	14

2.2.3 Resultados a la pregunta de investigación 1: ¿Existe un modelo de evaluación del HRI que incluya los factores humanos en entornos industriales?

Durante el proceso de revisión, no se ha encontrado ningún modelo de evaluación de la HRI ampliamente validado que incluya específicamente los factores humanos en entornos industriales. Sin embargo, se han identificado un total de 24 trabajos que llevan a cabo evaluaciones de HRI en contextos industriales. De estos trabajos, cinco de ellos se corresponden con marcos teóricos, otros cinco presentan herramientas específicas para la evaluación, y quince de ellos son estudios experimentales que han abordado la evaluación de la HRI en dichos entornos (Tabla 14). A pesar de la ausencia de un modelo consolidado, la presencia de estos estudios proporciona un panorama relevante sobre el interés y el desarrollo de investigaciones en el ámbito de la HRI en la industria.

Tabla 14: Clasificación de los artículos identificados por tipología.

Tipo de estudio	Nº de artículos	Referencias
Marco teórico	5	Cohen et al. (2018); Lindblom & Alenljung (2020); Meissner et al. (2020); Prati et al. (2021); Villani et al. (2018)
Herramientas	5	Charalambous et al. (2017); Gualtieri, Rauch, et al. (2020); Prati et al. (2021); Qbilat et al. (2021); von der Pütten & Bock (2018)
Estudios experimentales	15	Almeida et al. (2020); Aromaa et al. (2018); Baskaran et al. (2019); Beschi et al. (2020); Colim et al. (2021); Daniel et al. (2013); Danielsson et al. (2017); Harriott et al. (2013); Hietanen et al. (2020); Joosse et al. (2021); Kildal et al. (2019); Lasota & Shah (2015); Pantano et al. (2020); Schillaci et al. (2013); Tang & Webb (2018)

2.2.3.1 Marcos teóricos

A continuación, se presentan y analizan los cinco marcos teóricos identificados durante la revisión de la literatura. Estos marcos proporcionan una estructura conceptual y teórica para abordar la evaluación de la HRI en entornos industriales, permitiendo una mejor comprensión de los factores humanos involucrados y su impacto en el rendimiento y la experiencia de la persona usuaria. En esta sección se van a resumir los marcos teóricos identificados, que son:

- i) *The ANEMONE: Theoretical Foundations for UX Evaluation of Action and Intention Recognition in Human Robot Interaction* (Lindblom & Alenljung, 2020)
- ii) *Friend or Foe? Understanding Assembly Workers' Acceptance of Human-robot Collaboration* (Meissner et al., 2020)
- iii) *Survey on human-robot collaboration in industrial settings: Safety, intuitive interfaces and applications* (Villani et al., 2018)
- iv) *Workstation-Operator Interaction in 4.0 Era: WOI 4.0* (Cohen et al., 2018)
- v) *How to include User eXperience in the design of Human-Robot Interaction* (Prati et al., 2021)

2.2.3.1.1 *The ANEMONE: Theoretical Foundations for UX Evaluation of Action and Intention Recognition in Human-Robot Interaction* (Lindblom & Alenljung, 2020)

El enfoque de la metodología de evaluación ANEMONE se centra en (i) determinar si las personas pueden percibir, comprender y predecir las intenciones y acciones de un robot, y en qué medida, y (ii) en proporcionar una comprensión relevante de por qué algo funciona o no en una situación de uso y un contexto concretos (Lindblom & Alenljung, 2020).

Aunque ANEMONE sigue un procedimiento por fases, se considera un proceso iterativo en el cual puede ser necesario y ventajoso visitar algunas fases y subfases. Por ejemplo, la fase de fijación de objetivos podría requerir ajustes y objetivos más específicos después de un primer ciclo de evaluación exploratoria. El uso de esta metodología en un caso particular implica preparación, selección del tipo de evaluación, planificación y ejecución del procedimiento de evaluación, análisis e identificación de problemas. Estas fases se detallan en la Tabla 15:

Tabla 15: Descripción general del procedimiento de reconocimiento de acciones e intenciones en la HRI, fase por fase. Adaptado de Lindblom & Alenljung (2020).

Fase 1: Preparación
1.1 Identificación del caso: identificación del área de interés y definición del contexto
1.2 Especificar los objetivos de UX y de evaluación
Fase 2: Selección del tipo de evaluación de UX
2.1 Alcance de la evaluación
2.2 Caracterización de la evaluación analítica y empírica de la UX
Fase 3: Planificar y ejecutar la evaluación UX
3.1 Evaluación UX analítica del reconocimiento de acciones e intenciones
3.2 Evaluación UX empírica del reconocimiento de acciones e intenciones
Fase 4: Análisis de los datos recogidos e identificación de los problemas de UX
4.1 Estrategias para el análisis de los datos recogidos y la identificación de problemas
4.2 Análisis e identificación de problemas de reconocimiento de acciones e intenciones
Fase 5: Organización de los problemas de UX identificados en función del alcance y gravedad

El enfoque de ANEMONE se basa en el modelo de acción de siete etapas de Donald Norman (2013; 1986), el cual proporciona una estructura fundamental para analizar la interacción entre personas y objetos, incluida la interacción entre personas y robots industriales. Según este modelo, una acción se inicia como respuesta a un estado o evento en el contexto circundante, y se lleva a cabo para lograr un objetivo específico.

Según Norman, el ciclo de acción debe considerarse en su contexto de uso (D. Norman, 2014; D. A. Norman, 1991, 2005). La iniciación de la acción puede comenzar en la mente de una persona o como reacción a un estado de algo en el contexto circundante (objeto), es decir, una acción puede ser impulsada por un objetivo o por un evento (también llamado impulsado por datos), y se origina en el motivo general (actividad) de la persona usuaria (D. Norman, 2013; D. A. Norman, 1986, 1991, 2005). A continuación, se describen y se representan las etapas del modelo de Norman, modificadas por Lindblom & Alenljung (2020) incluyendo al robot como artefacto mediador (Figura 15):

- i) **Formación de la meta:** Existe una brecha entre el estado actual y el deseado que la persona quiere superar.
- ii) **Formación de la intención de actuar:** Se elabora un plan para hacer algo que supere la brecha y se esboza una dirección para lograrlo,
- iii) **Especificar la secuencia de acción:** Se formulan los detalles sobre cómo llevar a cabo la "estrategia" elegida.
- iv) **Ejecutar la acción:** Ejecutar físicamente la secuencia de acción especificada.
- v) **Percibir el estado del contexto (objeto):** Darse cuenta de las cosas que se pueden percibir.
- vi) **Interpretar el estado del contexto (objeto):** Concebir lo que significan las cosas percibidas.
- vii) **Evaluar el resultado en relación con los objetivos:** Comparar el nuevo estado con el deseado y valorar si el objetivo se ha cumplido en la medida suficiente.

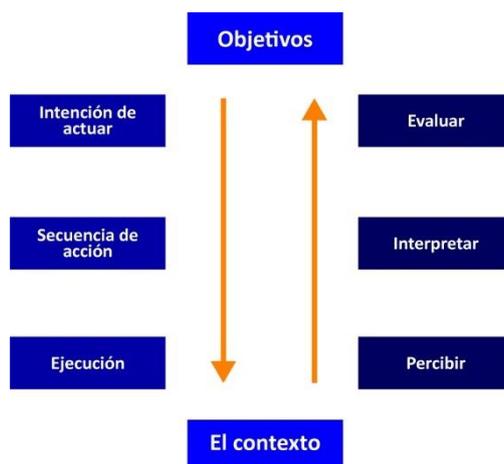


Figura 15: El modelo modificado de las siete etapas de la acción. Adaptado de Lindblom & Alenljung (2020).

Principales conclusiones

- El marco ANEMONE ofrece orientación sobre cómo medir, evaluar y valorar el reconocimiento mutuo de acciones e intenciones entre personas y robots. En particular, se centra en cómo identificar las posibilidades para que las personas perciban, comprendan y anticipen las intenciones y acciones de un robot, al tiempo que facilita la comprensión de por qué algo funciona o no en la interacción.

2.2.3.1.2 *Friend or Foe? Understanding Assembly Workers' Acceptance of Human-robot Collaboration* (Meissner et al., 2020)

Meissner et al. (2020) realizaron un estudio exploratorio con un enfoque teórico fundamentado en identificar los factores que probablemente influyan en la aceptación de las personas trabajadoras en contextos de HRC (Figura 16).

Tal y como puede apreciarse en la Figura 16 los factores primarios son los más influyentes (los que en la figura se encuentran en el centro). Estos, son el riesgo percibido (reducción de la plantilla, insuficiencia de la carga mental, exigencias del trabajo excesivas, aumento de la dependencia, disminución de la eficiencia y una menor interacción con los compañeros), los beneficios percibidos (alivio físico y mental, disminución de la tasa de error y el avance), las emociones negativas (como pueden ser la ansiedad, incertidumbre, impotencia, insuficiencia y frustración) y las emociones positivas (curiosidad y entusiasmo, diversión).

Además, los autores identifican una serie de factores secundarios: Factores relacionados con el objeto (seguridad, fiabilidad del proceso, adaptabilidad a las personas trabajadoras, adaptabilidad a la tarea y al entorno, diseño fiable y proceso controlable y autónomo), factores relacionados con el sujeto (posición del equipo, conocimientos y experiencias, autoevaluación, motivación laboral y miedo a perder el trabajo, confianza e interés en los robots, y personalidad y valores morales) y factores relacionados con el contexto (relación con los jefes, el proceso de implementación y relación con los compañeros).

Todo ello, influye en la actitud para la aceptación del sistema, además de que son factores estrechamente relacionados, y que, por lo tanto, también influyen entre sí.

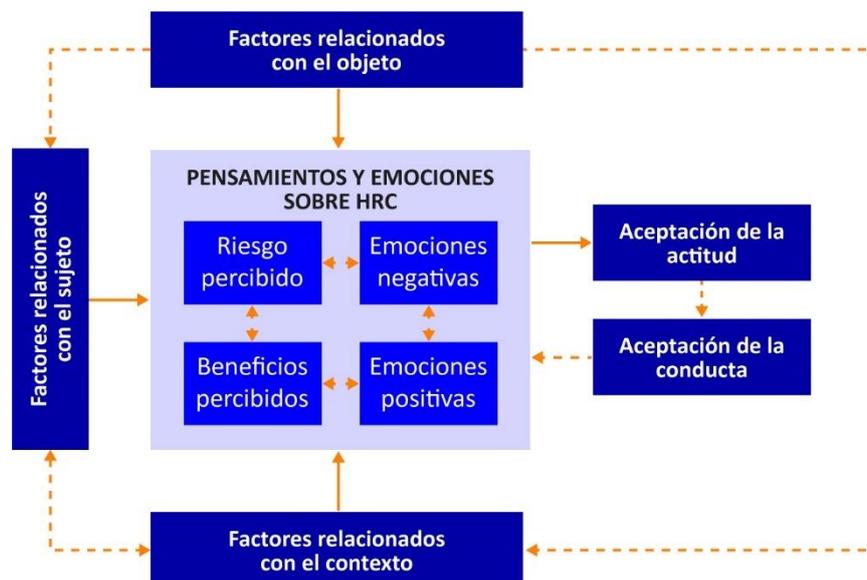


Figura 16: Factores identificados que influyen en la aceptación de la HRC. Adaptado de Meissner et al. (2020).

Principales conclusiones

- La aceptación de la HRC por parte de las personas trabajadoras de ensamblaje depende de su percepción de la HRC como una amenaza u oportunidad, la cual es influenciada por su evaluación de la tecnología (robots colaborativos) y el cambio organizativo asociado.
- Los pensamientos y sentimientos de las personas trabajadoras sobre la

HRC están influenciados por factores relacionados con el robot, el trasfondo individual de las personas trabajadoras y el entorno organizativo.

- La aceptación de la HRC no está determinada únicamente por consideraciones racionales, sino también por procesos cognitivos y emocionales.

2.2.3.1.3 *Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications* (Villani et al., 2018)

En el estudio de Villani et al. (2018) se presenta una amplia revisión de la colaboración persona-robot en un entorno industrial, con especial atención a las cuestiones relacionadas con la interacción física y cognitiva. Los autores identifican los principales retos de la robótica colaborativa en la industria, que incluyen la seguridad, el diseño de interfaces intuitivas y los métodos para optimizar la colaboración.

En cuanto a la seguridad, los autores destacan que es el principal desafío que afrontar en la HRI en entornos industriales. La colaboración entre personas y robots implica un contacto directo, y es crucial garantizar que esta interacción se realice de manera segura. Villani et al. (2018) proporcionan una clasificación de normas de seguridad relacionadas con la HRC.

Asimismo, enfatizan la importancia de diseñar interfaces de usuario intuitivas que permitan a las personas operarias interactuar fácilmente con los robots. Es esencial tener en cuenta aspectos relacionados con la seguridad y la salud mental de la persona operaria, tratando de evitar el estrés y la ansiedad derivados de la interacción con el robot. Para lograrlo, se deben diseñar interfaces que faciliten la comunicación y programación del robot de forma intuitiva, permitiendo que la persona operaria se enfoque en las tareas y objetivos. La información proporcionada por el robot debe ser la adecuada para que la persona usuaria sea consciente de la situación y pueda comprender el comportamiento del sistema y facilitar la intervención en situaciones dinámicas e imprevistas.

Villani et al. (2018) también resaltan la utilidad de las herramientas fisiológicas en el ámbito de la robótica afectiva, que busca mejorar la interacción al reconocer y adaptarse al afecto de la persona usuaria. La monitorización e interpretación de la comunicación no verbal puede proporcionar información importante sobre la interacción entre ambos y, por lo tanto, se puede conseguir una retroalimentación implícita sobre la interacción. En consecuencia, el objetivo de la robótica afectiva es aliviar la carga cognitiva de la persona usuaria cuando la tarea a realizar sobrecarga sus capacidades mentales, adaptando el comportamiento del robot e implementando un nivel suficiente de autonomía. Aunque estos enfoques se han centrado principalmente en robots de interacción social, podrían aplicarse en la robótica industrial para mejorar la comunicación y adaptación del robot según la carga cognitiva de la persona operaria. En consecuencia, el diseño de las interfaces persona-robot puede mejorarse teniendo en cuenta el procesamiento cognitivo de la información, la toma de decisiones, la percepción y otras capacidades o limitaciones de las personas.

El estudio también aborda la programación de los robots en aplicaciones industriales, donde se busca proporcionar instrucciones explícitas al robot de manera que resulte fácil y productivo para la persona operaria. Además, Villani et al. (2018) mencionan modos de interacción basados en la visión y la guía vocal, utilizados en la práctica industrial, y señalan la necesidad de abordar el reconocimiento del habla y el procesamiento del lenguaje para lograr una comunicación bidireccional natural. Concluyen que, cuando se considera el uso de interfaces de voz, hay que abordar dos aspectos: (i) el reconocimiento del habla, que implica el

reconocimiento de fonemas o palabras, y (ii) el procesamiento del lenguaje, que incluye la lectura y el análisis semántico. El objetivo final es establecer una comunicación bidireccional natural que permita que el robot entienda y genere el lenguaje natural. Para ello, proporcionar a las personas usuarias un *feedback* adecuado durante la interacción es una cuestión clave para el éxito de estos sistemas. En las aplicaciones industriales prácticas, una comunicación vocal basada en un lenguaje casi natural podría ser suficiente, en lugar del lenguaje natural, ya que el léxico que se debe utilizar es bastante limitado y las personas usuarias deben ser (al menos parcialmente) expertos en la interacción. Sin embargo, según Villani et al. (2018) la implementación y la validación de este enfoque se encuentra todavía a nivel de laboratorio.

Además, indican que, para abordar estos objetivos, es necesario abordar métodos de diseño adecuados, es decir, leyes de control, sensores y enfoques de planificación y asignación de tareas, que permitan a la persona estar cerca del robot de forma segura, compartiendo activamente el área de trabajo y las tareas, y proporcionando al sistema de interacción la flexibilidad necesaria. Para que la HRC sea eficaz en escenarios industriales reales es necesario la introducción de soluciones adaptativas para la robótica inclusiva. Destacan la necesidad de tener en cuenta a las personas usuarias vulnerables y, las diferentes habilidades y capacidades de las personas usuarias, así como en el diseño de soluciones colaborativas.

Villani et al. (2018) consideran que las aplicaciones industriales actuales de la HRC están en su mayoría infrautilizados, ya que se consideran principalmente herramientas que alivian la fatiga física de las personas trabajadoras y mejoran sus capacidades, pero gozan de una autonomía e inteligencia muy limitadas (Figura 17). Como objetivo futuro, señalan la necesidad de dotar a los robots de habilidades de procesamiento cognitivo adecuadas y de capacidades de autonomía compartidas, para que puedan asumir algunas tareas, aliviando así a las personas operarias del esfuerzo cognitivo, especialmente en tareas y escenarios complejos.



Figura 17: Visualización general de los resultados de Villani et al. (2018).

Principales conclusiones

- **La seguridad es el principal desafío** en la HRI en la industria. La colaboración directa entre personas y robots debe garantizarse de manera segura para evitar riesgos y accidentes.
- Es fundamental diseñar **interfaces de usuario intuitivas** que faciliten la interacción y programación del robot. Interfaces intuitivas permiten que las personas operarias se enfoquen en las tareas y objetivos, reduciendo el estrés y la ansiedad.
- La **incorporación de herramientas fisiológicas en la robótica afectiva** podría mejorar la interacción, reconociendo y adaptándose al afecto de la persona usuaria. Estos enfoques podrían aplicarse en la robótica industrial para una mejor comunicación y adaptación del robot según la carga cognitiva de la persona operaria.

- Los **enfoques basados en visión y guía vocal se utilizan en aplicaciones industriales para facilitar la interacción**. Reconocer el habla y procesar el lenguaje son aspectos cruciales para establecer una comunicación efectiva entre personas y robots.
- Para lograr una colaboración segura y flexible entre personas y robots en entornos industriales, **se deben aplicar métodos de diseño adecuados**, incluyendo leyes de control, sensores y enfoques de planificación.
- Para la colaboración persona-robot en entornos industriales, **es esencial tener en cuenta a las personas usuarias vulnerables y las diferentes habilidades y capacidades de las personas usuarias** al diseñar soluciones colaborativas.

2.2.3.1.4 Workstation–Operator Interaction in 4.0 Era: WOI 4.0 (Cohen et al., 2018)

El marco teórico presentado por Cohen et al. (2018) proporciona un enfoque integral para la interacción entre operarios y puestos de trabajo en la era 4.0. El marco se compone de tres elementos principales: Observación, Análisis y Reacción. Cada elemento recoge y procesa información relevante para lograr una comprensión global de la interacción entre la persona operaria y el entorno de trabajo (Figura 18).

El sistema de observación recopila datos de múltiples fuentes, incluyendo la persona operaria, el entorno del puesto de trabajo y las reacciones anteriores del sistema. Utiliza agentes de entrada y perfiles para capturar datos sobre el lenguaje corporal, gestos, movimientos, expresiones faciales, medidas fisiológicas y otras variables relevantes. El análisis de la información recopilada es esencial para comprender el estado de la persona operaria y del entorno de trabajo. Los datos se procesan para diagnosticar el estado fisiológico y cognitivo de la persona operaria, así como el estado del puesto de trabajo. Además, se clasifican los errores y se planifica la estrategia de reacción. El sistema selecciona el modo de funcionamiento más adecuado y la estrategia más apropiada en función de los datos actuales y las tendencias. La clasificación de estados y la planificación permiten tomar decisiones informadas sobre la mejor estrategia de reacción.

El elemento de Observación en el marco teórico propuesto por Cohen et al. (2018) adquiere información de tres fuentes fundamentales: la persona operaria, el entorno (incluido el puesto de trabajo) y el elemento de reacción, que responde a interacciones o eventos anteriores. Este proceso dinámico rastrea constantemente tanto al operador como al entorno circundante para inferir su estado actual, basándose en los cambios recientes. La implementación del marco contempla dos sistemas interconectados:

- a. **Sistema de Observación del Operario:** Este sistema está compuesto por diversos agentes que recopilan datos sobre el comportamiento del operario. Los agentes de entrada registran información relacionada con el lenguaje corporal, gestos, movimientos corporales, movimientos oculares, expresiones faciales, expresiones verbales y auditivas, así como medidas fisiológicas como frecuencia cardíaca y transpiración. Por otro lado, el agente de perfil recopila periódicamente los vectores de entrada y genera un nuevo vector que representa el estado granular estable de la persona operaria, basándose en información de segundos anteriores. Esta representación incluye indicadores fisiológicos (ritmo cardíaco, temperatura corporal, nivel de transpiración, etc.), posiciones y movimientos corporales (gestos de las manos, gestos del cuerpo, posición de piernas, etc.), actividad ocular (ritmo de parpadeo, ubicación del centro de visión, tamaño de la pupila, etc.), músculos faciales (postura de los labios, fruncimiento del ceño, etc.) y aspectos del lenguaje

(significado, volumen, tono, emoción, confianza, etc.). Por último, el agente de trayectoria del perfil integra periódicamente una secuencia de perfiles momentáneos y refleja las tendencias de cambio en ellos a lo largo de una unidad de tiempo mayor.

- b. **Sistema de Observación del Puesto de Trabajo:** Este sistema monitorea las condiciones de trabajo del operario, es decir, el estado del puesto de trabajo en sí. El agente de entrada recibe señales de datos continuas de diferentes sensores que registran aspectos físicos como el ruido, la temperatura, la luz, la humedad y los movimientos de fondo, transformándolos en un vector de valores que se actualiza de manera continua y frecuente. Además, el agente de trayectoria de perfiles registra los perfiles momentáneos de la estación de trabajo durante un periodo de tiempo más prolongado.

La función principal del sistema propuesto es llevar a cabo el análisis de la información recogida de forma continua del operario y del entorno. Este análisis tiene como objetivo obtener una comprensión global de lo que ocurre en el puesto de trabajo, incluyendo tanto el estado del operario como el estado del entorno. Además de construir un diagnóstico, el análisis busca seleccionar la mejor estrategia de reacción ante los estados específicos de un operador y un entorno de trabajo. El análisis del sistema comprende varios elementos interconectados, que son los siguientes:

- a. **Agente de análisis del estado de la persona operaria:** Este agente integra la información estática, momentánea y de tendencias del elemento de observación de la persona operaria para diagnosticar el estado general fisiológico y cognitivo de la persona operaria. Su función es clasificar el estado actual de la persona basándose en los datos recopilados.
- b. **Agente de análisis del estado del puesto de trabajo:** Similar al agente de análisis del estado de la persona operaria, este agente integra la información estática, momentánea y de tendencia del sistema de observación del puesto de trabajo para diagnosticar el estado del puesto de trabajo en sí.
- c. **Subsistema de análisis de errores:** Este subsistema interpreta y clasifica los datos recogidos durante la etapa de observación, identificando tanto la comisión de errores como secuencias de acciones que pueden conducir a errores. El subsistema despliega tres agentes: (i) prevención de errores, (ii) desarrollo y evaluación de la eficacia de la estrategia de reacción, y (iii) detección de la ocurrencia y rectificación de errores.
- d. **Agente de selección de modo:** El agente de selección de modo elige el modo de funcionamiento más adecuado en función de los datos actuales relacionados con los estados de la persona operaria y del puesto de trabajo, así como de las tendencias observadas en un repertorio de modos posibles.
- e. **Agente de decisión de la estrategia:** Este elemento se encarga de seleccionar la estrategia más adecuada en función del modo de operación y de los datos obtenidos sobre un operador específico.

El elemento de Reacción en el marco teórico de Cohen et al. (2018) se refiere a la aplicación de la estrategia de reacción determinada durante la etapa de análisis. El sistema traduce esta estrategia en una respuesta específica que mejor se ajuste a las circunstancias según los conocimientos obtenidos en el análisis. La implementación de la reacción se realiza considerando tres dimensiones:

- a. **El canal perceptivo:** Determina si se utilizará un mensaje sonoro, visual o táctil para comunicar la reacción al operario.
- b. **La frecuencia, el momento y la duración del mensaje:** Se ajustan según la situación para lograr una respuesta efectiva.

- c. **La intensidad:** Relacionada con el volumen, el brillo o la fuerza del mensaje, se adapta según el canal perceptivo empleado.

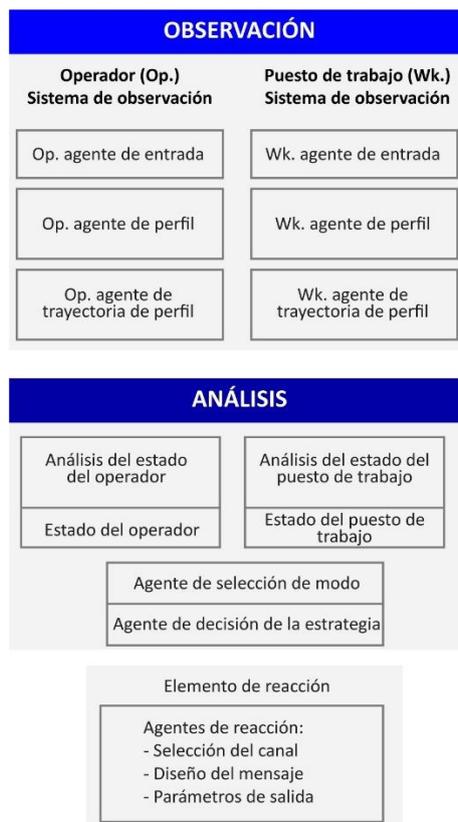


Figura 18: El marco de WOI 4.0 y su propuesta de arquitectura. Adaptado de Cohen et al. (2018).

Principales conclusiones

- El marco teórico proporciona un enfoque integral para la interacción entre operarios y puestos de trabajo en la era 4.0, a través de los elementos de Observación, Análisis y Reacción.
- El sistema de Observación recopila datos relevantes de múltiples fuentes, como la persona operaria, el entorno del puesto de trabajo y las reacciones anteriores, mediante agentes de entrada, perfiles y trayectorias.
- El Análisis es esencial para comprender el estado fisiológico y cognitivo de la persona operaria, así como el estado del puesto de trabajo. Permite clasificar errores y seleccionar la mejor estrategia de reacción basándose en los datos recogidos.
- La Reacción se traduce en una respuesta específica que mejor se adapte a las circunstancias, considerando las dimensiones del canal perceptivo, frecuencia, momento, duración e intensidad para comunicar la reacción al operario de manera efectiva.

2.2.3.1.5 *How to include User eXperience in the design of Human-Robot Interaction* (Prati et al., 2021)

Prati et al. (2021) proponen un método estructurado orientado a la UX para investigar el diálogo persona-robot, con el fin de mapear la interacción con los robots durante la ejecución de tareas compartidas y, finalmente, obtener los requisitos para el diseño de una HRI valiosa.

Las interfaces pueden generar diferentes tipos de comunicación: desde el lenguaje gráfico hasta la comunicación basada en la voz, pasando por el diálogo basado en los gestos. Además, los tipos de interfaces cambian en consecuencia. Por ejemplo, la comunicación gráfica puede tener lugar utilizando dispositivos específicos (por ejemplo, un monitor o una pantalla táctil), mientras que la comunicación basada en la voz puede utilizar interfaces de lenguaje natural y la comunicación basada en gestos puede utilizar cámaras adecuadas para seguir las manos del ser humano. Tal y como indican Prati et al. (2021), según la tipología de la comunicación, las interfaces persona-robot pueden clasificarse en cuatro categorías principales:

- i) Pantallas visuales (por ejemplo, interfaces gráficas de usuario, interfaces de realidad aumentada)
- ii) Gestos (por ejemplo, movimientos faciales y de la mano)
- iii) Habla y lenguaje natural (por ejemplo, respuestas auditivas y basadas en texto)
- iv) Interacciones físicas y hápticas

El sector industrial actualmente se inclina más por las interfaces visuales (Prati et al., 2021). Esto se debe, probablemente, a la multiplicidad de opciones de información que se ofrecen al robot y, en consecuencia, a la variedad de retroalimentación que se recibe del mismo. Además, simplemente ofrecen una representación clara e inequívoca del estado del trabajo al operario. Las interfaces de lenguaje natural se han investigado poco en el contexto industrial, mientras que ya se han aplicado para la interacción con robots sociales, como las expresiones faciales, el seguimiento de la mirada, la proxémica y la cinética, la háptica.

El enfoque propuesto por Prati et al. (2021) tiene como objetivo introducir un conjunto de técnicas de UX para apoyar el diseño de la interfaz persona-robot para construir aplicaciones HRI-HRC. De hecho, los autores destacan la necesidad de diseñar sistemáticamente una UX positiva. Según el enfoque centrado en las personas, las personas usuarias finales deben participar en todo el proceso de diseño, desde la primera fase de investigación hasta la evaluación final.

Por lo tanto, el primer paso para proceder a un enfoque UX es la creación de un equipo multidisciplinar. Este método puede garantizar un intercambio de información y una combinación de conocimientos de diferentes disciplinas para detallar el escenario de uso y tener una visión clara del resultado final.

A continuación, la persona usuaria se sitúa en el centro del proceso de diseño durante la llamada fase de "comprensión de la persona usuaria". Consiste en un análisis profundo y preciso de la persona usuaria, basado en la investigación del contexto y el análisis de la persona usuaria, que debe llevarse a cabo de forma más o menos invasiva. Las principales técnicas de UX adoptadas para la comprensión de la persona usuaria son: la observación de la persona usuaria, el grupo de discusión y la entrevista. Además, en el diseño de UX se pueden utilizar diferentes tipos de herramientas de mapeo para describir diversos aspectos de la interacción y los procesos asociados.

El flujo de trabajo basado en la UX se esquematiza en la Figura 19. De acuerdo con el enfoque propuesto, la investigación seleccionó un conjunto de herramientas específicas de UX

para apoyar el diseño de la interfaz persona-robot. Las herramientas adoptadas son las siguientes:

- i) Análisis de personas usuarias (observación, grupos focales, entrevista)
- ii) Análisis de tareas
- iii) Matriz usuario/tarea (nueva herramienta creada por los autores)
- iv) Mapas de experiencia (nueva herramienta creada por los autores)

Estas herramientas permiten llevar a cabo un sólido análisis de los requisitos de la persona usuaria centrado en las personas, que es la cuestión principal para garantizar un buen diseño de UX. Después, la fase de diseño definirá las principales características de los robots y del lugar de trabajo. El uso de prototipos es útil para comprobar si la solución identificada es la correcta. También en esta fase es posible involucrar a las personas usuarias para recoger sus opiniones y primeras impresiones. Esto facilitará la resolución de algunos problemas encontrados. Se ha definido un flujo de diseño de la HRI según un flujo de trabajo basado en la UX.

Una vez finalizado el diseño, las personas usuarias participan en la fase de evaluación del proyecto para comprobar la eficiencia, eficacia y satisfacción reales durante el uso del diseño final. En esta fase, la participación de las personas usuarias finales puede poner de manifiesto si hay aspectos que no habían sido considerados por el equipo de diseño.

Por lo tanto, el modelo se compone de los siguientes 4 pasos:

- i) Recogida de requisitos
 - i. Creación de un equipo multidisciplinar (por ejemplo, con la participación de ingenieros de sistemas, ingenieros informáticos, diseñadores de UI y de UX, desarrolladores de sistemas, integradores de sistemas)
 - ii. Análisis de personas usuarias (mediante observaciones, grupos de discusión y entrevistas)
 - iii. Análisis de actividades (mediante el análisis de tareas)
 - iv. Visualización de la interacción (mediante la matriz de personas usuarias/tareas y los mapas de experiencia)
- ii) Diseño de la interfaz
- iii) Creación de prototipos
- iv) Evaluación de la UX (basada en pruebas de personas usuarias).

Como todos los ciclos de diseño, el proceso puede no ser lineal, sino iterativo, y requiere muchos ciclos a través del proceso.

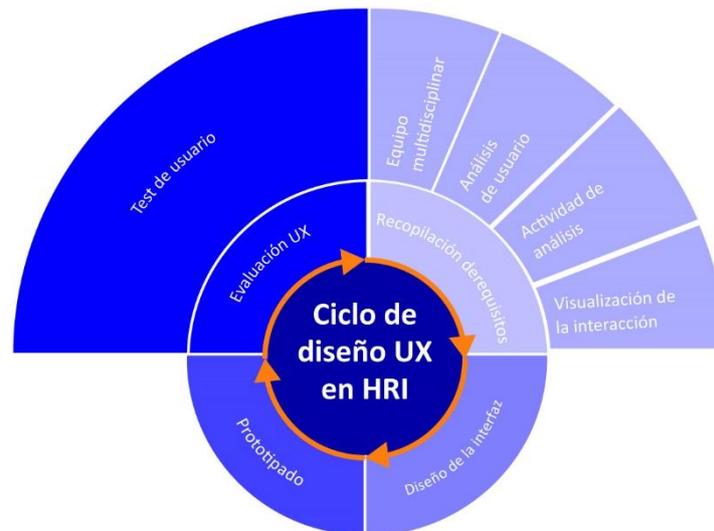


Figura 19: El flujo de trabajo de diseño UX propuesto para la HRI. Adaptado de Prati et al. (2021).

Principales conclusiones

- El flujo de diseño UX propuesto para la HRI es un método estructurado para investigar el diálogo persona-robot y para obtener requisitos para un diseño valioso de HRI. Destaca que incluye herramientas de diseño de interacción que analizan la UX, en lugar de centrarse únicamente en cuestiones tecnológicas.
- La investigación aplicó el método propuesto a un caso industrial centrado en operaciones de ensamblaje respaldadas por robots colaborativos y vehículos guiados automáticamente. Los resultados experimentales demostraron la validez del método propuesto para identificar necesidades de interacción y guiar el diseño de la interfaz.
- La investigación resalta la importancia de considerar la UX y los factores humanos en el diseño de HRI, especialmente en el contexto de la Industria 4.0 y la creciente colaboración entre robots y personas en fábricas inteligentes.

2.2.3.2 Herramientas identificadas

En el contexto de esta SLR, se han identificado diversas herramientas en los siguientes artículos, cada una de las cuales representa una contribución relevante en el ámbito de la HRI. Estas herramientas ofrecen enfoques y metodologías para abordar aspectos críticos relacionados con la UX, los factores humanos, la accesibilidad y la seguridad:

- The development of a Human Factors Readiness Level tool for implementing industrial human-robot collaboration* (Charalambous et al., 2017)
- Development and Validation of the Self-Efficacy in Human-Robot-Interaction Scale (SE-HRI)* (von der Pütten & Bock, 2018)
- Safety, Ergonomics and Efficiency in Human-Robot Collaborative Assembly: Design Guidelines and Requirements* (Gualtieri, Rauch, et al., 2020)
- A Proposal of Accessibility Guidelines for Human-Robot Interaction* (Qbilat et al., 2021)

v) *How to include User eXperience in the design of Human-Robot Interaction* (Prati et al., 2021)

2.2.3.2.1 *The development of a Human Factors Readiness Level tool for implementing industrial human-robot collaboration* (Charalambous et al., 2017)

En el estudio de Charalambous et al. (2017) propusieron un sistema para conocer el nivel de madurez industrial para desarrollar una nueva herramienta de nivel de preparación de las personas para que los profesionales del diseño de sistemas optimicen la implementación satisfactoria de la HRC industrial (Figura 20).

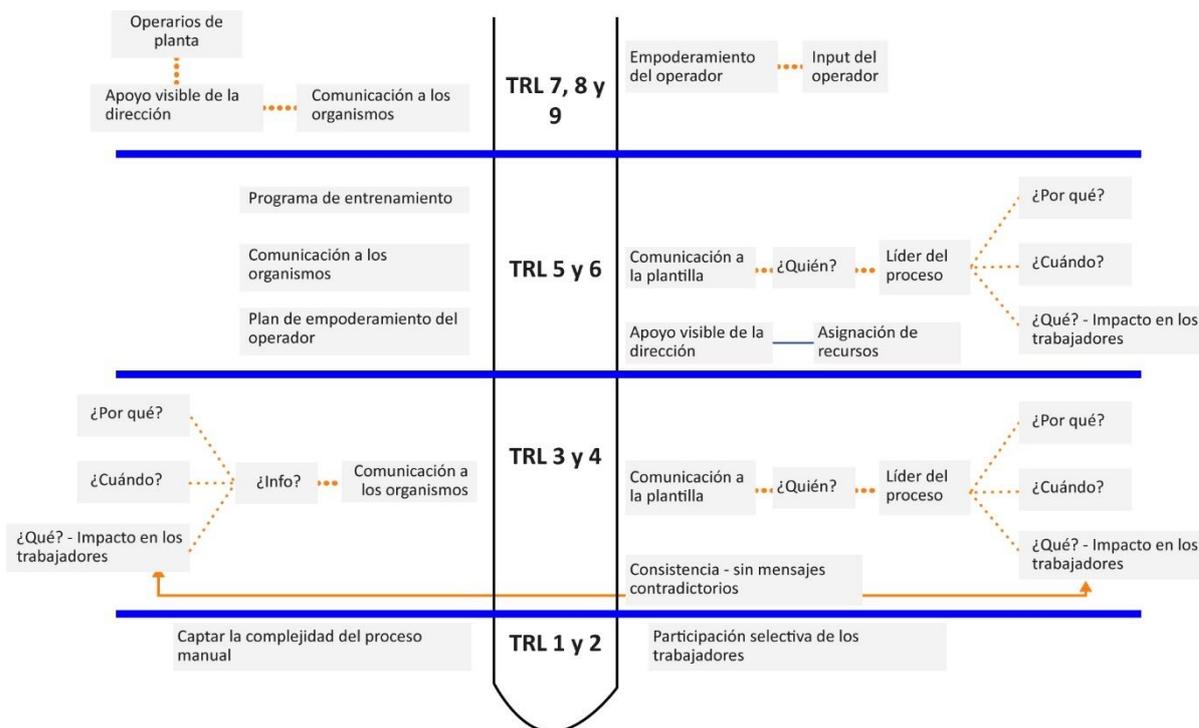


Figura 20: Esquema para los profesionales que indica los factores humanos organizativos clave en diferentes TRL. Adaptado de Charalambous et al. (2017).

2.2.3.2.2 *Development and Validation of the Self-Efficacy in Human-Robot-Interaction Scale (SE-HRI)* (von der Pütten & Bock, 2018)

En el estudio de von der Pütten & Bock (2018), se desarrolla y valida una nueva medida de autoeficacia en la HRI. A través de varios estudios experimentales, proponen un cuestionario final que consta de de 18 ítems (Tabla 16). Las personas participantes deben valorar los ítems en una escala de Likert de 6 puntos, donde 1 representa "Totalmente en desacuerdo" y 6 representa "Totalmente de acuerdo".

Este cuestionario ha sido validado mediante una muestra amplia de participantes alemanes e ingleses. Su objetivo es evaluar las percepciones de autoeficacia de las personas con respecto a la HRI. Además, los autores señalan que también puede utilizarse para evaluar la eficacia de programas de formación en este ámbito.

Tabla 16: Ítems del cuestionario SE-HRI. Adaptado de von der Pütten & Bock (2018).

Nº	Ítem
1	<i>I could set up a robot according to my wishes and my environment.</i>
2	<i>I could get a robot to perform a specific task.</i>
3	<i>I am familiar with technology; therefore, I think I could use a robot.</i>
4	<i>I think I could adjust a robot the way that it could help me in my daily life.</i>
5	<i>It is easy to use a robot.</i>
6	<i>If I should solve a problem with the assistance of a robot, I could do that.</i>
7	<i>To achieve a specific goal with the assistance of a robot will not be a problem for me.</i>
8	<i>I could teach a robot something if I would try hard enough.</i>
9	<i>I could easily learn how to use a robot.</i>
10	<i>I could teach a robot to complete easy tasks.</i>
11	<i>If I would use a robot, I would always know how and why it behaves like it does.</i>
12	<i>I could do easy adjustments on a robot by myself.</i>
13	<i>I could use a robot in daily life.</i>
14	<i>I would feel comfortable while interacting with the robot.</i>
15	<i>If a robot is doing something wrong, I could find a way to change its behavior.</i>
16	<i>Robots are easy to control.</i>
17	<i>I could deploy a robot in a specific way to save time.</i>
18	<i>I am very confident in my abilities to control a robot.</i>

2.2.3.2.3 *Safety, Ergonomics and Efficiency in Human-Robot Collaborative Assembly: Design Guidelines and Requirements* (Gultieri, Rauch, et al., 2020)

En el estudio de Gultieri, Rauch, et al. (2020) se desarrolla una recopilación y clasificación de requisitos previos y directrices de diseño. Estas directrices de diseño podrían ayudar a los diseñadores a desarrollar y evaluar adecuadamente *workstations* colaborativos de ensamblaje seguras, centradas en las personas y eficientes.

En esta guía, se tienen en cuenta los siguientes requisitos:

- i) Requisitos de seguridad
- ii) Requisitos de ergonomía física
- iii) Requisitos relacionados con la ergonomía cognitiva
- iv) Requisitos relacionados con la eficiencia

2.2.3.2.4 *A Proposal of Accessibility Guidelines for Human-Robot Interaction* (Qbilat et al., 2021)

Qbilat et al. (2021) proponen pautas de accesibilidad para la HRI. Estas pautas fueron evaluadas por diecisiete diseñadores y desarrolladores de HRI a través de un cuestionario que constaba de nueve preguntas en una escala Likert de cinco puntos. Además, se incluyeron seis preguntas abiertas para evaluar las pautas propuestas desde cuatro perspectivas principales: usabilidad, aceptación social, experiencia del usuario y su impacto social. En la Tabla 17 se presenta la clasificación de las pautas de accesibilidad.

Tabla 17: Clasificación general de los requisitos de las directrices propuestas. Adaptado de Qbilat et al. (2021)

Aspecto	Punto de control
Perceptible	Múltiples modalidades de interacción Color y contraste Ubicación de los componentes de <i>hardware</i> y <i>software</i> Alternativas para elementos no textuales Componentes parpadeantes Contenido visual intermitente Pantallas Tecnología de asistencia e interfaces web
Manejable	Controles de <i>hardware</i> y funcionamiento físico Teclas, teclados y mandos Navegación en las pantallas Tiempo
Entendible	Interacción predecible Errores, ayuda y retroalimentación Voz natural Muestra
General	Adopción de las preferencias de interacción de la persona usuaria Asistencia humana accesible

2.2.3.2.5 *How to include User eXperience in the design of Human-Robot Interaction* (Prati et al., 2021)

En el estudio realizado por Prati et al. (2021), además de presentar un marco teórico, se proponen dos herramientas de diseño para la HRI:

i) User/Task Matrix

Esta herramienta se utiliza para sintetizar en una tabla toda la información relevante sobre las personas usuarias, las tareas y las condiciones operativas (Figura 16). La matriz considera las tareas a realizar, su secuencia y las relaciones temporales entre ellas. También asigna quién realizará las diferentes acciones, ya sea el operario o los robots. Además, describe las condiciones específicas de trabajo, como el entorno (por ejemplo, ruidoso, pequeño, polvoriento) y el uso de equipos personales (por ejemplo, guantes, gafas, casco). Esta información sobre los actores implicados y las situaciones en las que se producen errores y cómo resolverlos resulta útil para diseñar la interfaz. La plantilla de la herramienta se muestra en la Figura 16, donde las filas enumeran las tareas y las columnas tienen en cuenta los elementos mencionados anteriormente. La matriz usuario/tarea ayuda a los diseñadores a definir las interfaces más adecuadas y debe ser actualizada periódicamente a medida que avanza el proyecto y el escenario de interacción se vuelve más definido.

ii) Experience Maps

Estos mapas permiten representar una visualización sintética de toda la experiencia que atraviesa un usuario "genérico" para lograr un objetivo específico. Se utilizan para comprender el comportamiento general de las personas en lugar de centrarse en aspectos específicos de un negocio particular, como lo hacen los mapas de viaje. Los Mapas de Experiencia se centran en la interacción de la persona operaria y mapean su "viaje" a lo largo del proceso. Esta herramienta se refiere a la experiencia de interacción entre la persona usuaria y el entorno que lo rodea y está estrechamente relacionada con las actividades realizadas. No incluye emociones ni sentimientos en esta etapa, ya que son difíciles de predecir y dependen de personas usuarias específicas, que se desconocen en esta fase del

proyecto. Sin embargo, se considera la posibilidad de agregar referencias al estado de ánimo de la persona usuaria en etapas posteriores.

2.2.3.3 Estudios experimentales

Además, la SLR ha permitido identificar 15 estudios experimentales, que son:

- *Evaluating the Effect of Saliency Detection and Attention Manipulation in Human-Robot Interaction* (Schillaci et al., 2013)
- *Simplified Human-Robot Interaction: Modeling and Evaluation* (Daniel et al., 2013)
- *Assessing physical workload for human-robot peer-based teams* (Harriott et al., 2013)
- *Analyzing the Effects of Human-Aware Motion Planning on Human-Robot Collaboration* (Lasota & Shah, 2015)
- *Assessing Instructions in Augmented Reality for Human-robot Collaborative Assembly by Using Demonstrators* (Danielsson et al., 2017)
- *Digital Human Models in Human Factors and Ergonomics Evaluation of Gesture Interfaces* (Aromaa et al., 2018)
- *The Design and Evaluation of an Ergonomic Contactless Gesture Control System for Industrial Robots* (Tang & Webb, 2018)
- *Empowering assembly workers with cognitive disabilities by working with collaborative robots: a study to capture design requirements* (Kildal et al., 2019)
- *Digital Human and Robot Simulation in Automotive Assembly using Siemens Process Simulate: A Feasibility Study* (Baskaran et al., 2019)
- *Interface Transparency Issues in Teleoperation* (Almeida et al., 2020)
- *AR-based interaction for human-robot collaborative manufacturing* (Hietanen et al., 2020)
- *A human-cyber-physical system approach to lean automation using an industrie 4.0 reference architecture* (Pantano et al., 2020)
- *How motion planning affects human factors in human-robot collaboration* (Beschi et al., 2020)
- *Lean Manufacturing and Ergonomics Integration: Defining Productivity and Wellbeing Indicators in a Human-Robot Workstation* (Colim et al., 2021)
- *Making Appearances: How Robots Should Approach People* (Joosse et al., 2021)

2.2.3.3.1 *Evaluating the Effect of Saliency Detection and Attention Manipulation in Human-Robot Interaction* (Schillaci et al., 2013)

El objetivo del estudio de Schillaci et al. (2013) era identificar las características físicas y de comportamiento que deben destacarse para garantizar una buena UX en la HRI, y medir la UX al interactuar con un robot equipado con mecanismos de atención. Para ello, adoptaron dos técnicas: (i) la UX medida mediante preguntas cualitativas y cuantitativas en cuestionarios; y, (ii) la proxémica, estimada a partir de los vídeos grabados durante las interacciones.

El robot utilizado durante la experimentación es el robot NAO, un robot humanoide de 57 cm de alto. Probaron la experimentación en cuatro combinaciones del sistema de atención del robot: (i) en la exploración, el robot se siente atraído por los movimientos, las caras y los objetos, y parece que explora el entorno; (ii) el comportamiento Interacción el robot mira y señala un objeto; (iii) Evitar la interacción, reproduce la pérdida de interés y el aburrimiento (el robot mira lejos del objeto entregado por la persona usuaria); y, (iv) Interacción completa, se compone de una secuencia de los comportamientos anteriores, la primera acción realizada es

la exploración, una vez el robot ha detectado una persona con la que interactuar y un objeto que puede utilizarse para llamar la atención de la persona usuaria, su estado pasa a ser de interacción, y tras un cierto periodo, pasa a evitar la interacción, a lo que sigue la exploración.

La experimentación consistía en que el robot realizara dichos comportamientos en cuatro sesiones de interacción distintas, uno por cada uno de los comportamientos. El participante tenía que sentarse frente al robot, mientras grababan la sesión. Después de cada una de las cuatro sesiones, se pedía a las personas participantes que rellenaran un cuestionario.

El cuestionario que utilizaron es el Godspeed (Bartneck et al., 2008), un cuestionario que utiliza pares semánticos para evaluar la actitud de las personas hacia el robot. El cuestionario contiene preguntas sobre cinco conceptos: Antropomorfismo, Animación, Simpatía, Inteligencia y Seguridad percibidas.

Los autores concluyen que el cuestionario utilizado es un buen instrumento para medir la calidad de los comportamientos de los robots implementados. Las correlaciones positivas entre los atributos antropomórficos y la inteligencia percibida confirman que un robot con apariencia humana puede aumentar el nivel de su inteligencia percibida. Sin embargo, una apariencia excesivamente humana puede suponer que la persona que interactúa tenga unas expectativas demasiado altas sobre las capacidades cognitivas del robot, lo que puede provocar una decepción cuando el robot no cumpla dichas expectativas.

Además, los autores indican que las personas participantes no intentaron comunicarse vocalmente con el robot, lo que sugiere que no esperaban esta modalidad de interacción debido a la ausencia de una boca en la cara del robot y a la capacidad verbal de cualquier otro robot.

Los autores concluyen que una contribución relevante en la satisfacción de la persona usuaria viene dada por la capacidad de respuesta y la interactividad del robot y puede incrementarse mejorando su sistema de retroalimentación. Consideran que proporcionar información visual y auditiva al participante es de suma importancia para aumentar la intuición de la interacción y la satisfacción de las personas usuarias.

Los autores concluyeron que el cuestionario Godspeed resultó ser una herramienta eficaz para medir la calidad de los comportamientos implementados en el robot. Además, se encontraron correlaciones positivas entre los atributos antropomórficos y la percepción de inteligencia del robot, lo que sugiere que una apariencia más humana puede aumentar la percepción de su inteligencia. Sin embargo, una apariencia excesivamente humana puede generar expectativas poco realistas sobre las capacidades cognitivas del robot, lo que podría llevar a una decepción si no se cumplen esas expectativas. Las personas participantes no intentaron comunicarse vocalmente con el robot, lo que indica que no esperaban esta modalidad de interacción debido a la ausencia de una boca en la cara del robot y a la capacidad verbal de otros robots.

2.2.3.3.2 *Simplified Human-Robot Interaction: Modeling and Evaluation* (Daniel et al., 2013)

El estudio de Daniel et al. (2013) tenía como objetivo mejorar la HRI mediante la implementación de una UI. Para ello, realizaron un experimento de laboratorio con el fin de verificar la mejora del rendimiento del enfoque Operación de Robots Orientada al Servicio (SORO, por sus siglas en inglés). El experimento constó de dos fases: (i) la persona usuaria tenía que ejecutar dos tareas con el robot a través de una interfaz tradicional (TGUI) y (ii) la persona usuaria tenía que ejecutar las dos mismas tareas a través de una interfaz flexible (FGUI).

Durante la ejecución de las tareas realizaron grabaciones de audio y vídeo. Además, se recogió el tiempo necesario para ejecutar la tarea, el número de interacciones realizadas y la ratio entre las interacciones con la pantalla táctil y el teclado.

Se administraron cuatro cuestionarios a las personas participantes:

- i) Encuesta sociodemográfica previa al test donde recogían datos como la edad, el sexo, educación y conocimientos en tecnología, mecánica, informática y robótica.
- ii) Se formuló una breve serie de preguntas sobre la confianza en la automatización en general.
- iii) Después de cada parte del experimento, se pidió al participante que respondiera a las preguntas de la entrevista (adaptadas de WAI Site Usability Testing Questions) sobre las impresiones sobre la UI actual y el funcionamiento de la célula robótica.
- iv) Se rellenó un Cuestionario de Condiciones de Flujo y se comprobó la confianza en el sistema actual.

Tras analizar los datos obtenidos de la experimentación, los autores determinaron que las interfaces de usuario que utilizan el enfoque SORO en cuanto a la creación de interfaces de usuario sencillas y claras facilitan el uso de los robots industriales y reducen significativamente el tiempo de configuración y funcionamiento. Además, demuestran que el enfoque SORO y el diseño cuidadoso de las interfaces de usuario permiten crear una HRI simplificada. Su ventaja es que el funcionamiento de la célula robótica depende menos de los conocimientos y la experiencia previa de la persona usuaria.

Finalmente, concluyeron que la UI simplificada es una solución efectiva para mostrar información relevante al operador. Además, es destacable la propuesta de un nuevo concepto de modelado de HRI que incluye factores como la confianza en la automatización, la conciencia situacional, la experiencia y las expectativas

2.2.3.3.3 *Assessing physical workload for human-robot peer-based teams* (Harriott et al., 2013)

El estudio de Harriott et al. (2013) se enfoca en el análisis de las métricas relacionadas con la carga de trabajo físico, la evaluación de modelos predictivos de carga de trabajo físico y la exploración del impacto de las situaciones de trabajo en equipo entre personas y robots en la carga de trabajo físico. En este estudio se presentan dos evaluaciones:

- i) La evaluación guiada: Implicó que las personas participantes siguieran instrucciones proporcionadas por un compañero, ya sea una persona ubicada en un lugar remoto o un robot local. Durante esta evaluación, las personas participantes debían evaluar el estado de "víctimas" (maniqués) mientras seguían las indicaciones. En algunos casos, las indicaciones eran proporcionadas por una persona y en otros por un robot. Tras completar la evaluación, las personas participantes verbalizaron sus percepciones subjetivas de carga de trabajo y luego completaron el cuestionario NASA-TLX.
- ii) La evaluación colaborativa: Requirió que las personas participantes trabajaran en colaboración y en equipo tomando decisiones conjuntas con un compañero humano o un robot local. Las personas participantes fueron informadas de la existencia de una amenaza de bomba anónima y se les encomendó la tarea de buscar objetos fuera de lugar y comunicar sus hallazgos a su compañero (que podía ser una persona o un robot). Después de investigar las áreas indicadas, las personas participantes completaron el cuestionario NASA-TLX.

La evaluación colaborativa requirió que las personas participantes trabajaran en colaboración y en equipo tomando decisiones conjuntas con un compañero humano o un robot local. Las personas participantes fueron informadas de la existencia de una amenaza de bomba anónima y se les encomendó la tarea de buscar objetos fuera de lugar y comunicar sus hallazgos a su compañero (que podía ser una persona o un robot). Después de investigar las áreas indicadas, las personas participantes completaron el cuestionario NASA-TLX.

Los autores concluyeron que los análisis de la carga de trabajo global en equipos humano-humano (H-H) y humano-robot (H-R) demostraron que el equipo H-R generó significativamente una carga de trabajo inferior en las evaluaciones Guiada y Colaborativa. A pesar de que se observó un mayor movimiento en las condiciones H-R en ambas evaluaciones, se llevó a cabo una investigación específica sobre el impacto de la colaboración en la carga de trabajo físico.

Las métricas subjetivas en ambas evaluaciones indicaron que las personas participantes en la condición H-R calificaron subjetivamente su carga de trabajo físico igual o inferior a las personas participantes en la condición H-H.

2.2.3.3.4 *Analyzing the Effects of Human-Aware Motion Planning on Human-Robot Collaboration* (Lasota & Shah, 2015)

El estudio de Lasota & Shah (2015) aborda la necesidad de que los robots colaboren con las personas en espacios de trabajo compartidos para aumentar la productividad y eficiencia. Destaca la importancia de permitir que los robots naveguen en espacios compartidos y el desarrollo de algoritmos y marcos de planificación de trayectorias especializados que consideren el factor humano. También proporciona una visión general de la investigación realizada en apoyo de este objetivo en diversos dominios complementarios.

El objetivo del estudio realizado por Lasota & Shah (2015) fue examinar la respuesta humana a la adaptación del robot a nivel de planificación de movimiento para determinar su efecto en la fluidez del equipo, la satisfacción humana y la percepción de seguridad y comodidad. Para ello, llevaron a cabo un experimento en el que las personas participantes colaboraron con un robot para realizar una tarea conjunta. El robot utilizado en el experimento fue el ABB IRB-120, un robot industrial estándar con 6 grados de libertad. Se utilizó un sistema de captura de movimiento Phase Space para rastrear las acciones realizadas por el ser humano en el espacio de trabajo. Además, se implementó un sistema de seguridad en tiempo real en el robot para ajustar su velocidad en función de la distancia de separación con la persona.

El experimento consistió en trabajar con un robot adaptable que incorporaba la planificación de movimientos con conciencia humana y con un robot de referencia que utilizaba movimientos estándar más cortos. Se evaluó la fluidez del equipo mediante métricas cuantitativas, mientras que la satisfacción humana y la percepción de seguridad y comodidad se evaluaron mediante cuestionarios. Los resultados demostraron que cuando las personas trabajaban con el robot adaptable, completaban la tarea de manera más rápida (un 5,57% más), con más movimientos concurrentes (un 19,9% más), menos tiempo de inactividad humano (un 2,96% menos), menos tiempo de inactividad del robot (un 17,3% menos) y una mayor distancia de separación (un 15,1% más) en comparación con el robot estándar. Además, las respuestas al cuestionario indicaron que las personas participantes se sentían más seguros y cómodos trabajando con el robot adaptable, y estaban más satisfechos con él como compañero de equipo en comparación con el robot estándar.

Lasota & Shah (2015) determinaron que las personas responden favorablemente a la adaptación del robot a nivel de planificación de movimiento, lo cual puede proporcionar importantes beneficios tanto en la fluidez de la interacción entre personas y robots como en la

satisfacción de la persona trabajadora. Asimismo, destacan que la planificación de movimientos con conciencia humana puede aumentar la percepción de seguridad y reducir el riesgo de problemas de salud relacionados con el estrés, a la vez que mejora la fluidez del equipo.

Los autores concluyeron que la necesidad de la investigación se basó en la importancia de permitir que los robots colaboren con personas en espacios compartidos para aumentar la productividad y eficiencia, destacando la relevancia de desarrollar algoritmos que consideren el factor humano. La metodología empleada en el estudio consistió en un experimento con participantes personas, en el cual se calcularon el tiempo de ejecución de la tarea, el movimiento concurrente, la distancia promedio de separación, el tiempo inactivo del robot y el tiempo inactivo de la persona. Además, las personas participantes completaron un cuestionario para evaluar la satisfacción y seguridad percibida. Los resultados muestran que la planificación de movimiento consciente del robot condujo a una interacción más eficiente y satisfactoria entre personas y robots, mejorando la fluidez del equipo y la satisfacción de la persona trabajadora.

2.2.3.3.5 *Assessing Instructions in Augmented Reality for Human-robot Collaborative Assembly by Using Demonstrators* (Danielsson et al., 2017)

Danielsson et al. (2017) abordan el tema de la colaboración estrecha entre personas y robots en la industria manufacturera, con el objetivo de mejorar la eficiencia y accesibilidad en el proceso de fabricación. Los autores señalan que virtualizar las instrucciones y la interacción entre operadores y robots puede ser una solución más adaptable y accesible para las personas trabajadoras.

El estudio se enfoca en dos objetivos principales: evaluar la capacidad de los demostradores para simular la HRC y determinar si las interfaces basadas en RA pueden guiar a las personas en pruebas de ensamblaje previamente desconocido. A largo plazo, el propósito del demostrador es funcionar como un módulo de prueba para instruir de manera eficiente a las personas operarias que colaboran con robots en la industria manufacturera.

La metodología empleada en este trabajo involucra dos iteraciones de pruebas del demostrador con voluntarios en dos exposiciones diferentes. Los grupos de prueba consistieron en cuatro grupos de estudiantes de secundaria de escuelas técnicas, considerados futura mano de obra de la industria manufacturera, lo que los hace representativos para este contexto. La primera iteración fue un estudio piloto para evaluar la viabilidad de la tarea de ensamblaje y la intuición de la interfaz del demostrador. La segunda iteración se desarrolló como un estudio de usuarios, utilizando un enfoque de métodos mixtos que combinó datos cuantitativos, como los tiempos de finalización de la tarea de ensamblaje, y datos cualitativos, obtenidos a través de cuestionarios de usabilidad.

En cada grupo, se eligieron tres personas voluntarias para realizar el montaje utilizando el demostrador, mientras que las otras dos observaban. Los voluntarios fueron seleccionados de aquellos que levantaron primero la mano para participar y, en general, mostraron una predisposición positiva hacia la adopción de nuevas tecnologías. Aunque los grupos eran mixtos, los voluntarios eran todos hombres. Luego de que cada persona completara la tarea, se procedía con la siguiente persona y así sucesivamente.

Después de cada prueba, se solicitó a las personas de prueba que completaran un cuestionario de usabilidad, concretamente el cuestionario SUS, mientras que los estudiantes observadores también respondieron un cuestionario que evaluaba su interés general y la información mostrada en la pantalla del demostrador durante la prueba.

Los resultados obtenidos mostraron que el demostrador fue factible e intuitivo para la mayoría de las personas de prueba, sin embargo, todos cometieron errores en al menos uno

de los pasos del ensamblaje y la cantidad de errores al trabajar de forma independiente fue inaceptablemente alta. Esto indicó que las instrucciones proporcionadas no fueron lo suficientemente claras. Además, los datos recopilados del estudio de usuarios, representados mediante las puntuaciones SUS, presentaron una amplia variabilidad entre los grupos, lo que dificultó obtener conclusiones concluyentes. El alineamiento entre el mundo virtual y el real no fue perfecto, lo que afectó la comprensibilidad de las instrucciones para las personas de prueba. Se identificó que la traducción del cuestionario para las personas de prueba del inglés al sueco también pudo haber afectado los resultados.

Un resumen del protocolo de prueba muestra que de las 12 personas que participaron en la prueba, todas cometieron errores en al menos uno de los pasos. De un total de 144 pasos de montaje, 75 pasos se realizaron con al menos una desviación de las instrucciones dadas. En las discusiones abiertas después de las pruebas, muchas de las personas que participaron en la prueba y los estudiantes del grupo de observación señalaron que no estaba claro que dónde debían leer las instrucciones.

Los autores concluyen que el demostrador ha alcanzado un nivel de madurez que permite a personas sin experiencia previa en el montaje trabajar de forma independiente en todos los pasos del demostrador, lo que responde al segundo objetivo de este trabajo. Sin embargo, la cantidad de errores al trabajar de forma independiente es demasiado alta para ser aceptable. La cantidad de errores muestra que el montaje es lo suficientemente complejo como para requerir instrucciones.

Como conclusión, el uso de demostradores y RA tiene potencial en la HRC en la industria manufacturera. La virtualización de instrucciones y la interacción mediante interfaces basadas en RA ofrecen una solución adaptable y accesible para mejorar la eficiencia en el proceso de fabricación. El demostrador ha alcanzado un nivel de madurez que permite a personas sin experiencia previa en ensamblaje trabajar de forma independiente en todos los pasos del proceso. Esto sugiere que los demostradores pueden ser herramientas útiles para simular la HRC en tareas de ensamblaje complejas. Aunque el demostrador fue factible e intuitivo para la mayoría de las personas de prueba, se observó un alto número de errores al trabajar de forma independiente. La cantidad de errores indica la necesidad de mejorar la claridad y precisión de las instrucciones proporcionadas, lo que es crucial para facilitar la colaboración efectiva entre personas y robots en la industria manufacturera.

2.2.3.3.6 *Digital Human Models in Human Factors and Ergonomics Evaluation of Gesture Interfaces* (Aromaa et al., 2018)

El propósito del estudio llevado a cabo por Aromaa et al. (2018) fue examinar si la utilización de Modelos Humanos Digitales (DHM, por sus siglas en inglés) por parte de expertos en Factores Humanos/Ergonomía podría enriquecer la evaluación de prototipos de interfaces gestuales. Con tal fin, llevaron a cabo dos casos de estudio en los cuales se evaluaron sistemas basados en gestos para el control remoto de robots.

En el primer caso de estudio, la tarea de la persona usuaria consistía en operar un sistema robótico a distancia mediante gestos realizados con la mano derecha. El objetivo era emplear el sistema robótico para transportar un objeto pequeño desde el punto A al punto B. La tarea abarcaba desde conducir el vehículo hasta el objeto, mover el brazo robótico a una posición específica, tomar el objeto con la pinza del brazo robótico, dirigir el sistema hacia una caja y finalmente soltar el objeto dentro de dicha caja.

El segundo caso de estudio replicó el procedimiento del primero, pero el sistema basado en visión por computadora adoptaba un enfoque absoluto y utilizaba la articulación del codo para generar gestos diferentes.

Según los autores, los casos de estudio ilustraron la importancia de considerar tanto los aspectos mentales como físicos de las personas al diseñar interfaces gestuales, además de las métricas de desempeño convencionales (como el tiempo y los errores). Además, destacaron que la incorporación del cuestionario NASA-TLX permitió identificar mejoras que quizás no hubiesen salido a la luz solo a través del rendimiento.

Aromaa et al. (2018) concluyeron que los resultados indicaron que las interfaces gestuales podrían incrementar la carga física sobre las personas usuarias. Por ende, subrayan la relevancia de abordar ampliamente los aspectos de factores humanos/ergonomía durante el diseño de vocabularios gestuales, en vez de enfocarse exclusivamente en las medidas de tiempo y error. La evaluación de los trastornos musculoesqueléticos emerge como un asunto crucial, especialmente al diseñar aplicaciones destinadas a contextos laborales industriales.

Basándose en este estudio, se sugiere que la aplicación de DHM podría enriquecer la evaluación de factores humanos/ergonomía. Los autores resaltan que los análisis posturales realizados mediante los DHM concuerdan con la experiencia de fatiga en la muñeca y el brazo. Además, sugieren que los DHM resultan particularmente adecuados para evaluar interfaces gestuales que involucran el uso de grupos musculares más amplios.

2.2.3.3.7 *The Design and Evaluation of an Ergonomic Contactless Gesture Control System for Industrial Robots* (Tang & Webb, 2018)

En el estudio llevado a cabo por Tang y Webb (2018), se aborda el diseño y evaluación de un sistema de control gestual sin contacto ergonómico para robots industriales en entornos colaborativos persona-robot. Tradicionalmente, la manipulación de robots se realiza mediante consolas de aprendizaje, pero en contextos colaborativos se demanda una comunicación más flexible e intuitiva. El control gestual omnipresente ofrece esta versatilidad al permitir a las personas usuarias interactuar con los robots sin dispositivos físicos, reduciendo la carga física y favoreciendo la cooperación fluida.

El estudio realizó evaluaciones para abordar problemas ergonómicos en el diseño y comparar la UX de este sistema con un controlador de aprendizaje convencional. Las personas participantes realizaron pruebas en dos condiciones: control gestual y controlador de aprendizaje. Los criterios ergonómicos evaluados fueron la carga física, la intuitividad y el grado de disfrute.

Antes de la experimentación, las personas participantes rellenaron un cuestionario demográfico (edad, sexo y cargo profesional). Seguido, recibieron un resumen de cinco minutos sobre los fundamentos del robot industrial. Antes de la prueba de cada interfaz, las personas participantes tuvieron 5 minutos de periodo de familiarización y entrenamiento. El tiempo de prueba de cada sistema duró diez minutos. Al acabar con cada una de las interfaces, las personas participantes rellenaron el cuestionario SUS (Brooke, 1996). Pidieron a las personas participantes que puntuaran el sistema en función de tres criterios ergonómicos que son: (i) la carga de trabajo físico, (ii) la intuitividad y (iii) el grado de disfrute.

Como la experimentación se centra en la evaluación de los aspectos ergonómicos de los sistemas, las personas participantes no tenían que realizar ninguna tarea específica. Se les indicó que primero debían probar todos los movimientos posibles con cada sistema y luego manipular la pinza robótica alrededor de los tubos de plástico sin golpearlos. Se les dijo que cada prueba terminaba en diez minutos, pero que podían parar en cualquier momento.

Obtuvieron como resultado una mayor puntuación en la evaluación rápida de las extremidades superiores (RULA, por sus siglas en inglés) para la interfaz de mando de aprendizaje, lo que indica mayor riesgo de lesiones musculoesqueléticas tras un uso prolongado. En el cuestionario la interfaz de gestos obtuvo mejor puntuación. Sin embargo, los autores indican que algunos participantes señalaron que la información visual de la interfaz

gráfica de usuario y la información táctil de la pantalla táctil han facilitado su uso, lo que indica el beneficio potencial de incorporar información visual y táctil al sistema de control de gestos. Aunque, la mitad de las personas participantes señalaron que el mando de aprendizaje es pesado de sostener.

Los autores concluyen que los experimentos demostraron que el sistema de control por gestos desarrollado tiene potencial para ser utilizado como dispositivo de entrada para el control de robots industriales en un escenario de colaboración persona-robot. Sin embargo, en el futuro podrían incorporarse más funciones para mejorar la facilidad de uso y la intuición. Indican que esta podría mejorarse si se complementa con una interfaz gráfica con elementos convencionales que proporcionen al usuario un *feedback* visual.

2.2.3.3.8 *Empowering assembly workers with cognitive disabilities by working with collaborative robots: a study to capture design requirements* (Kildal et al., 2019)

Kildal et al. (2019) destacan la importancia de abordar los desafíos que enfrentan las empresas manufactureras en los entornos de mercado actuales, como la reducción de los tiempos de desarrollo de productos, la adaptación a ciclos de vida más cortos, la creciente demanda de personalización y la competencia global. Para hacer frente a estos desafíos, se propone una nueva metodología que analiza la arquitectura funcional y física de productos existentes para identificar familias de productos orientadas al ensamblaje.

En el estudio realizado por Kildal et al. (2019), se llevó a cabo una evaluación piloto cualitativa en la que un trabajador de montaje con discapacidades cognitivas y expertos multidisciplinarios interactuaron con un robot colaborativo en tareas de montaje rutinarias. Las personas expertas involucradas en el estudio incluyeron técnicos de apoyo en la línea de montaje, personas responsables de la adaptación de requisitos e instrucciones de montaje, médicos expertos en ergonomía, psicólogos clínicos y técnicos expertos en prevención de riesgos laborales. La metodología del estudio piloto consistió en llevar a cabo una evaluación de experto con personas usuarias finales en diferentes roles de partes interesadas, y con expertos en temas relevantes. En concreto, los siguientes expertos participaron en el estudio:

- Técnico de apoyo en la línea de montaje con experiencia en la prestación de apoyo a personas con discapacidad cognitiva que realizan trabajos de montaje.
- Persona responsable de la adaptación de los requisitos e instrucciones de montaje del cliente final.
- Doctor en medicina, experto en ergonomía, con amplia experiencia en la adaptación del puesto de trabajo para personas con discapacidades cognitivas.
- Psicólogo clínico.
- Técnico experto en prevención de riesgos laborales, especializado en trabajadores con discapacidad.

Además, un trabajador con discapacidades cognitivas también participó en la evaluación del sistema prototipo. La persona trabajadora tenía experiencia en la realización de montajes de armarios eléctricos de forma manual en la cadena de montaje de la organización objetivo. Las personas expertas observaron cómo la persona trabajadora desempeñaba las tareas encomendadas junto con el robot.

El resultado de la evaluación piloto fue la identificación de directrices generales para el diseño de sistemas de HRI en el contexto del montaje de productos. Estas directrices incluyeron la utilización de canales de comunicación multimodales redundantes para permitir que las personas trabajadoras elijan dónde centrar su atención, la personalización de las formas en que las personas trabajadoras se comunican con el robot y la importancia de no

permitir que las personas trabajadoras pierdan habilidades de montaje al depender excesivamente del robot.

- i) Utilizar canales de comunicación multimodales redundantes, para que las personas trabajadoras puedan elegir dónde centrar su atención y puedan recibir información de estado importante incluso cuando su atención sensorial esté cambiando o si no se acuerdan de comprobar pantallas específicas.
- ii) Permitir la personalización de las formas en las que un trabajador se comunica con el robot, con opciones como no utilizar la salida de voz si así se solicita.
- iii) Tener siempre en cuenta que las personas trabajadoras no deben perder habilidades de montaje porque se sientan obligados a tener que confiar en el robot para hacer cosas que ya saben hacer (es decir e., ofrecer la asistencia mínima necesaria para que las personas trabajadoras puedan acceder a trabajos que hasta entonces no eran accesibles para ellos).

2.2.3.3.9 *Digital Human and Robot Simulation in Automotive Assembly using Siemens Process Simulate: A Feasibility Study* (Baskaran et al., 2019)

El objetivo del estudio de Baskaran et al. (2019) es utilizar una herramienta de fabricación digital para simular un proceso de ensamblaje de automóviles físicamente exigente y utilizar un modelo humano para analizar el impacto ergonómico. Los autores, llevaron a cabo diferentes pruebas de simulación para modelos humanos (masculinos y femeninos). A continuación, utilizaron una evaluación combinada de la colaboración entre personas y robots.

La simulación del proceso de ensamblaje se realizó con *Siemens Process Simulate V14.0*. A través de esta plataforma, se puede evaluar la interacción entre las personas que trabajan en la línea de montaje, los equipos y el flujo de materiales. A efectos de simulación, sólo se consideró la barra de torsión de mayor complejidad geométrica y peso, por lo que se necesitó un único gemelo digital de la persona para realizar las tareas. El modelo gemelo digital humano se crea a partir de las tablas estandarizadas de características biomecánicas, antropométricas y ergonómicas mediante el *software Process Simulate*. Una vez creado el entorno de simulación, el siguiente paso consistió en seleccionar entre diferentes modelos humanos para realizar la operación de montaje.

Los autores concluyen que la simulación virtual proporciona al diseñador la libertad de probar el proceso de ensamblaje con diferentes soluciones potenciales y le permite crear un proceso óptimo en términos de tiempo e impacto ergonómico en las personas. Los autores resumen sus principales conclusiones en estos puntos:

- Los informes ergonómicos generados a partir de diferentes ensayos de simulación realizados con éxito (con peso y altura variables) no mostraron diferencias significativas entre las cuatro métricas ergonómicas y fatiga en los escenarios exclusivamente humanos.
- Basado en el análisis de los informes ergonómicos generados a partir de diferentes ensayos de simulación con altura y peso variables, se pudo inferir que los modelos humanos femeninos con alturas inferiores a 1630 mm no son capaces de colocar la barra de torsión en la orientación deseada y completar la operación. Esta limitación se superó posteriormente introduciendo una solución robótica colaborativa para llevar la barra de torsión a la ubicación correcta y sujetarla para la persona trabajadora.

2.2.3.3.10 *Interface Transparency Issues in Teleoperation* (Almeida et al., 2020)

El objetivo de la experimentación de Almeida et al. (2020) era comprender la influencia de una interfaz de visualización inmersiva en relación con la percepción espacial y del

movimiento. Para ello, los autores diseñaron varias tareas en las que las personas usuarias tenían que indicar una posición 3D en el espacio basándose en la información visual mediada por medios tecnológicos: un monitor con una sola vista amplia de la escena remota (interfaz tradicional) frente a múltiples vistas parciales de la escena adquiridas de forma natural.

Para analizar el efecto de la retroalimentación visual inmersiva, se pidió a las personas participantes que realizaran tres tareas de coordinación mano-ojo utilizando dos interfaces visuales: la interfaz tradicional (monitor fijo, un único punto de vista con una escena de visión amplia) frente a la interfaz inmersiva (HMD y punto de vista controlable). Las personas participantes empezaron al azar, con la interfaz 1 o 2, para mitigar el efecto del factor de aprendizaje.

La evaluación consistió en el análisis de un conjunto de parámetros relacionados con el rendimiento, que se recogieron durante los experimentos, y de las respuestas dadas a un breve cuestionario después de cada prueba. La evaluación de la usabilidad la realizaron en dos partes: (i) la evaluación objetiva, basada en el desempeño de la tarea y, (ii) la evaluación subjetiva, crearon un cuestionario inspirado en el IBM Computer Usability Satisfaction Questionnaire (J. R. Lewis, 2009) y basado en las preguntas de presencia de Slater et al. (1994) y Usoh et al. (2000).

2.2.3.3.11 *AR-based interaction for human-robot collaborative manufacturing* (Hietanen et al., 2020)

Hietanen et al. (2020) proponen un modelo basado en sensores de profundidad para la monitorización del espacio de trabajo y una UI interactiva de RA para una HRC segura. La UI de RA se implementa en dos equipos diferentes: un proyector-espejo y un equipo de RA portátil. El modelo de espacio de trabajo y las UI se evalúan en una tarea realista de montaje de motores diésel.

Las personas participantes tenían que realizar una tarea que, a su vez, se dividía en cinco sub-tareas:

- i) Tarea 1: Instalar 8 ejes de balancín (H)
- ii) Tarea 2: Instalar el bastidor del motor (R)
- iii) Tarea 3: Insertar 4 tornillos del bastidor (H)
- iv) Tarea 4: Instalar el eje de balancines (R+H)
- v) Tarea 5: Introducir las tuercas en el eje (H)

Durante la ejecución de las tareas, se recogieron el tiempo de desempeño. Después de la ejecución de las tareas, las personas participantes rellenaron un cuestionario para evaluar el estrés físico y mental.

Los resultados obtenidos indican que las UI interactivas basadas en la RA proporcionan una reducción del 21-24% y del 57-64% en el tiempo de finalización de la tarea y del tiempo de inactividad del robot, respectivamente, en comparación con una línea de base sin interacción y espacio de trabajo compartido. Sin embargo, la evaluación de la UX revela que la RA basada en HoloLens aún no es adecuada para la fabricación industrial, mientras que la configuración del proyector-espejo muestra claras mejoras en la seguridad y la ergonomía del trabajo.

2.2.3.3.12 *A human-cyber-physical system approach to lean automation using an industrie 4.0 reference architecture* (Pantano et al., 2020)

En el estudio de Pantano et al. (2020) proponen una estructura de diseño que aborda la seguridad, las interfaces y los métodos de diseño. Su experimentación debía evaluar si la introducción de las técnicas era suficiente. Crearon un procedimiento para simular los tres riesgos. En el caso de la colisión y el aplastamiento, los peligros se simulaban entrando en la zona peligrosa y obstruyendo el movimiento del robot.

Se evaluó la UI, más concretamente, la carga de trabajo percibida por las personas usuarias al relacionarse con el robot y su interfaz. Para proceder a la evaluación, identificaron una tarea de colaboración con el robot. En la tarea, la persona usuaria tenía que recuperar una maqueta de una pieza de fabricación y medir las dimensiones de dos puntos. Después de la medición, tenían que escribir las medidas en una hoja de papel. Para medir cómo afectaba la interfase al usuario, la tarea debía realizarse con tres medios de interacción. El primero era el medio de interacción normal proporcionado por el robot a través de un *smartpad*. El segundo fue la interfaz de nuevo diseño con entrada sólo táctil. El tercero fue la interfaz de nueva creación con entradas para la mirada y el tacto. Las personas participantes rellenaron un cuestionario que se componía de NASA-TLX y de una pregunta de tres opciones en la que se preguntaba cuál era la preferencia de la persona usuaria.

2.2.3.3.13 *How motion planning affects human factors in human-robot collaboration*
(Beschi et al., 2020)

El estudio de Beschi et al. (2020) recoge dos experimentaciones. Los objetivos de la primera experimentación eran (i) comprender la influencia de los cambios de trayectoria en el comportamiento humano en los espacios de trabajo, (ii) comprender la influencia de los cambios de velocidad y aceleración en el comportamiento humano, y (iii) comprender si el ser humano se sincroniza con el robot. Para ello, se pidió a las personas participantes que ejecutaran una serie de tareas que no estaban sincronizadas con el robot.

La segunda experimentación requería sincronización entre el robot y la persona. Los objetivos de la segunda experimentación eran (i) entender cómo la persona operaria se ve influenciado por los cambios de velocidad, aceleración y tiempo de ejecución, y (ii) entender cómo la persona operaria percibe la cooperación persona-robot cuando tiene que sincronizar sus actividades con el robot.

En ambas experimentaciones tuvieron en cuenta medidas subjetivas. Para ello utilizaron un cuestionario, en la primera experimentación se trataba de un cuestionario que evalúa el nivel de riesgo e interferencia percibido por la persona. En la segunda experimentación el cuestionario utilizado evaluaba si las personas participantes habían notado cambios en la configuración de la planificación de la movilidad y con qué configuración se sienten más cómodos.

En la primera experimentación utilizaron también medidas objetivas como el tiempo necesario, que dio lugar al cálculo de la ratio de tiempo requerido con y sin el robot. También calcularon la distancia entre la persona y el robot.

Beschi et al. (2020) indican que sus resultados manifiestan que el cambio de trayectoria influye en la atención de la persona operaria con respecto a los cambios de velocidad/aceleración. Por el contrario, el rendimiento humano se ve influido por el rendimiento del robot. El ser humano también estima las distancias del robot, especialmente durante los movimientos predecibles. La trayectoria del ser humano no vio influenciada por los movimientos del robot, ya que la trayectoria del robot no intercepta la trayectoria de la persona operaria.

En cuanto a las emociones de la persona operaria, el tiempo de trayectoria es más importante que la velocidad y las aceleraciones. Los diferentes perfiles de movimiento con el mismo tiempo total son difíciles de reconocer cuando la persona operaria centra su atención en la tarea.

2.2.3.3.14 *Lean Manufacturing and Ergonomics Integration: Defining Productivity and Wellbeing Indicators in a Human–Robot Workstation* (Colim et al., 2021)

El estudio realizado por Colim et al. (2021) se centró en objetivos principales: (i) evaluar el riesgo musculoesquelético asociado al puesto de premontaje antes y después de la implementación del control HRC; (ii) evaluar el riesgo musculoesquelético vinculado al puesto de pre-ensamblado antes y después de la incorporación del cobot, mediante un enfoque multimétodo; y, (iii) analizar los indicadores de bienestar y aceptación de la robótica, basándose en las percepciones de las personas trabajadoras.

Para alcanzar estos objetivos, consideraron varios indicadores clave:

En cuanto al desempeño y análisis de residuos, evaluaron (i) el tiempo de ejecución, (ii) la variabilidad (que se refiere a las fluctuaciones en los tiempos de producción) y (iii) el índice de producción. El enfoque principal fue minimizar o eliminar las tareas que no agregaran valor al proceso de producción. A través de la observación directa, identificaron las tareas de premontaje como uno de tres tipos: tareas que agregan valor, tareas que no agregan valor y tareas que no agregan valor, pero son necesarias. Estas últimas no tienen un impacto directo en el proceso, pero son esenciales para la creación del producto final y, por lo tanto, deben reducirse en la medida de lo posible.

En lo que respecta al riesgo musculoesquelético, llevaron a cabo una evaluación multimétodo que incluyó la RULA, el Índice de Tensión Revisado (RSI) y el Método de Indicadores Clave para evaluar la carga de trabajo físico durante las operaciones de manipulación manual.

Además, diseñaron un cuestionario para recopilar las percepciones de las personas trabajadoras sobre la robótica en la industria y las mejoras ergonómicas asociadas. Este cuestionario abordó aspectos como la caracterización personal de las personas trabajadoras, el impacto de la robótica en su entorno laboral, el esfuerzo percibido en las tareas y la evaluación general de su puesto de trabajo.

Colim et al. (2021) concluyeron que será fundamental desarrollar metodologías para la evaluación de la ergonomía cognitiva de los sistemas HRC. La muestra limitada de trabajadores participantes es la principal limitación de este estudio. En general, los resultados del presente estudio demuestran que la integración de LM y E&HF en una estación de trabajo HRC potencia la implantación exitosa de esta tecnología y la mejora continua de los procesos de fabricación. Por lo tanto, este enfoque multimétodo puede utilizarse en el desarrollo e implementación de entornos de la Industria 4.0, siendo el factor humano el punto central.

2.2.3.3.15 *Making Appearances: How Robots Should Approach People* (Joose et al., 2021)

En el estudio de Joosse et al. (2021) realizan dos experimentaciones con el fin de verificar si sus hipótesis son ciertas o no.

En la primera experimentación analizaron las reacciones de las personas participantes ante una persona o un robot que se acercaba a ellos más de lo que generalmente se considera apropiado, es decir, que invadía su espacio personal. En esta experimentación pretenden validar dos hipótesis:

- i) Hipótesis 1: La gente responde más negativamente cuando una persona invade su espacio personal que cuando lo hace un robot.
- ii) Hipótesis 2: Las personas mostrarán más comportamientos compensatorios cuando se les acerque un robot a mayor velocidad, en comparación con cuando una persona se acerque a mayor velocidad.

Diseñaron una tarea en la que las personas participantes permanecían de pie en una posición para permitir que los agentes se acercaran a ellos. Sin embargo, no se lo mencionaban explícitamente a las personas participantes de antemano. Durante la tarea, las personas participantes debían encontrar figuras en un cartel de "Dónde está Waldo". Para evitar que las personas participantes completaran la tarea antes de ser abordados por el agente, dos de los cinco personajes fueron eliminados digitalmente del cartel y, por tanto, no pudieron ser encontrados.

El procedimiento de la experimentación fue el siguiente: El participante debía cumplimentar un formulario de consentimiento, después el participante era llevado a la sala de experimentación, donde ya se encontraba el agente (que en algunos casos fue una persona y en otros el robot humanoide Nomad Scout). Seguido, el participante comenzaba a buscar las figuras en el cartel. Un minuto después de que el participante comenzara a buscar las figuras, el agente iniciaba la aproximación. Una vez que se producía la invasión del espacio personal, el experimentador conducía a las personas participantes a la sala de información, donde se les pidió que rellenaran un cuestionario. La duración de cada sesión fue de aproximadamente 20 minutos.

El cuestionario que debían rellenar medía las actitudes hacia los robots, el atractivo físico y social, el parecido humano y la confianza (más detallado en el estudio de Joosse et al. (Joosse et al., 2013)). Además, la experimentación se grabó con dos cámaras. Para codificar los datos de vídeo se utilizaron mediciones conductuales (objetivas). Los datos se grabaron con dos cámaras, una encima del rotafolio donde estaba colgado el póster para grabar las expresiones faciales y la postura de la cabeza, y otra en la esquina más alejada de la sala, que ofrecía una visión general para grabar los movimientos de todo el cuerpo, como alejarse. Los codificadores codificaron los comportamientos identificados por la investigación sociopsicológica como comportamientos (observables) que se producen cuando se invade el espacio personal.

Para la primera hipótesis, predijeron que las personas responderían más negativamente y, por tanto, mostrarían más comportamientos compensatorios cuando una persona invadiera su espacio personal en comparación con un robot. Cuando consideraron la suma de los comportamientos compensatorios (sumando los comportamientos emocionales positivos, los emocionales negativos y los de evitación física), comprobaron que las personas participantes mostraban efectivamente más comportamientos compensatorios cuando se les acercaba una persona que un robot.

Para la segunda hipótesis, predijeron que una velocidad de aproximación más rápida conduciría a un comportamiento compensatorio mayor cuando el robot invadiera el espacio personal de una persona. Sin embargo, obtuvieron como resultado que la velocidad no influyó en los comportamientos compensatorios cuando se trataba de la persona. Sin embargo, una mayor velocidad de aproximación del robot condujo a un comportamiento compensatorio emocional menos negativo, aunque este efecto no fue significativo. Por lo tanto, los autores rechazan la segunda hipótesis.

La segunda experimentación tenía como objetivo validar las siguientes cuatro hipótesis:

- iii) Hipótesis 3: Un robot que se acerca con un ruido intencionado será percibido más positivamente que un robot que se acerca con un ruido no intencionado.
- iv) Hipótesis 4: Un robot que se aproxima con una velocidad de aproximación intencionada será percibido más positivamente que un robot sin velocidad de aproximación intencionada.

- v) Hipótesis 5: Un robot que se acerca con patrones de ruido y velocidad congruentes será percibido más positivamente que un robot que se acerca con patrones de ruido y velocidad incongruentes.
- vi) Hipótesis 6: (a) La gente percibirá un robot más alto como más inteligente que un robot más bajo, y, (b) la gente se sentirá más segura cuando se le acerque un robot más bajo que un robot más alto.

Por lo tanto, para validar las hipótesis mencionadas, tienen en cuenta las siguientes variables independientes y realizaron una experimentación 2 x 2 x 2:

Ruido: Crearon dos patrones de ruido diferentes para transmitir la intención (de velocidad) del robot variando el nivel de ruido. La intensidad del ruido en el primer patrón variaba, utilizando así el ruido de forma funcional para transmitir información sobre las intenciones de movimiento del robot. El segundo patrón no variaba en intensidad, por lo que no proporcionaba ninguna pista sobre la intención de velocidad del robot; considerándolo "ruido no intencionado".

Velocidad: Los robots se programaron para acelerar y desacelerar, bien lentamente en el tiempo y conducir de forma "suave", o bien para acelerar y desacelerar lo más rápido posible, y conducir de forma "brusca". Al primer patrón de aceleración lo denominaron "velocidad intencionada", con el fin de comunicar la intención de acercarse y detenerse. El segundo patrón de aceleración tenía una velocidad más constante, por lo que se denominó "velocidad no intencionada".

Las dos variables mencionadas dieron lugar a cuatro condiciones experimentales para cada uno de los dos robots.

Altura del robot: Cada participante experimentó el acercamiento de uno de los dos robots: un Magabot o un Giraff. El Giraff es un robot de telepresencia de 163 cm de altura, mientras que el Magabot mide 78 cm. En general, su aspecto se asemejaba a un ordenador con ruedas y ambos tenían un diseño bastante funcional. En las pantallas de ambos robots, se mostraban un par de ojos, formados por puntos de color estáticos.

Las personas participantes primeramente firmaron un formulario de consentimiento. A continuación, se pidió a las personas participantes que se colocaran en un punto marcado en el suelo. No se les exigió que se situaran exactamente en el centro del lugar, ni se les prohibió explícitamente que se movieran cuando el robot se acercaba. Cuando las personas participantes estaban en su sitio, el experimentador inició un guion que hacía que el robot se dirigiera en línea recta hacia la posición del participante. Las aproximaciones se grabaron con tres cámaras: dos cámaras situadas en la esquina del laboratorio del experimento y orientadas hacia las personas participantes (una con un primer plano y otra con un ángulo más amplio) y una cámara de ojo de pez colocada en el techo para grabar la escena general. Después de que el robot completara su aproximación, las personas participantes debían completar el cuestionario Godspeed (Bartneck et al., 2008).

Las condiciones de ruido intencionado recibieron puntuaciones más altas que las condiciones de ruido no intencionado, lo que los autores interpretan como una evaluación más positiva del robot en estas condiciones concretas, apoyando así la tercera hipótesis.

Los autores no encontraron efectos para la velocidad de aproximación en ninguna de las escalas de Godspeed (Bartneck et al., 2008), por lo que rechazan la cuarta hipótesis, que afirmaba que una aproximación con velocidad de aproximación intencionada sería percibida más positivamente que un robot que se acercara sin velocidad de aproximación intencionada.

Encontraron una diferencia significativa en la percepción de la velocidad cuando era congruente con el ruido (tanto el ruido intencionado como la velocidad, o tanto el ruido no intencionado como la velocidad), en comparación con cuando el comportamiento no era

congruente (ruido intencionado con velocidad no intencionada, o ruido no intencionado con velocidad intencionada). En las condiciones congruentes, se percibió que los robots se acercaban a una velocidad adecuada, mientras que, en las condiciones incongruentes, se percibió que ambos robots se acercaban demasiado rápido. No hubo diferencias significativas en las demás variables dependientes, por lo que no se acepta la hipótesis 5.

La altura del robot influyó en la seguridad percibida. Las personas participantes percibieron el Magabot corto como significativamente más seguro que el Giraff más alto apoyando la hipótesis 6b. Ninguna de las demás variables dependientes se vio influida por el robot concreto utilizado. Dado que el robot más alto no fue percibido como más inteligente, rechazaron la hipótesis 6b.

2.2.3.3.16 Visión general de los estudios experimentales

A continuación, se ofrece una visión general de los estudios experimentales identificados. Se exploran tres aspectos de estos estudios: (i) el tamaño de la muestra y la representación del género en las personas participantes, (ii) los tipos de robots utilizados en las investigaciones y, (iii) un resumen de los resultados y enfoques empleados en los estudios experimentales.

- *Tamaño de la muestra y perspectiva de género*

Se ha recogido el tamaño de la muestra utilizada en los diferentes casos de estudio y el número de hombre y mujeres en cada uno de ellos. En la Tabla 18 se muestra cuántas personas han participado en cada uno de los casos de estudio, de los cuales cuántos eran hombres y cuántas mujeres. Se observa, que tan solo el 30% de la muestra total corresponde a mujeres, y que el 65% corresponde a hombres. Siendo un total de 450 participantes, tan solo 134 eran mujeres frente a 291 hombres.

Tabla 18: Tamaño de la muestra de los estudios experimentales.

Ref.	Número de personas	Mujeres	Hombres	Otros
Schillaci et al. (2013)	28	8	20	0
Daniel et al. (2013)	16	4	12	0
Lasota & Shah (2015)	20	7	13	0
Danielsson et al. (2017)	25	7	17	1
	27	7	19	1
	25	7	18	0
	21	4	14	1
Aromaa et al. (2018)	10	2	8	0
	9	4	5	0
Tang & Webb (2018)	8	4	4	0
Almeida et al. (2020)	25	2	23	0
Hietanen et al. (2020)	20	nd	nd	nd
Pantano et al. (2020)	19	4	15	0
Beschi et al (2020)	32	4	28	0
Joose et al. (2021)	85	35	50	0
	80	35	45	0
Total	450	134	291	3

▪ *Tipos de robots utilizados en los estudios experimentales*

Durante la fase de ejecución, la mayoría de los estudios han empleado diversos tipos de robots, siendo los brazos robóticos los más comunes. En la Tabla 19 se detalla el robot utilizado en cada experimentación. No obstante, en el estudio de Baskaran et al. (2019), la evaluación se realizó mediante el *software* Siemens Process, por lo que los autores prescindieron por completo de la utilización de robots físicos. De manera similar, Colim et al. (2021) llevaron a cabo la experimentación en una estación de trabajo sin emplear ningún robot físico, al igual que en el estudio de Almeida et al. (2020), que se centró en las interfaces.

Tabla 19: Tipo de robot utilizado en cada estudio experimental.

Tipo de robot	Robot	Ref.
Robot humanoide	Robot NAO	Schillaci et al. (2013)
Brazo robótico	Nachi MC70	Daniel et al. (2013)
	ABB IRB-120	Lasota & Shah (2015)
	YumiIRB 14000	Kildal et al. (2019)
	Universal RobotsUR5	Tang & Webb (2018) y Hietanen et al. (2020)
	UniversalRobotsUR10	Beschi et al. (2020)
	KUKA LBR IIWA 14 R820	Pantano et al. (2020)
Robot integral	Un robot que consiste en: (1) un vehículo todo terreno de tracción en las cuatro ruedas controlado a distancia, (2) un brazo robótico Lynx al5d y (3) una cámara con movimiento panorámico e inclinación.	Aromaa et al. (2018)
	Nomad Scout	Joose et al. (2021)
	Magabot	Joose et al. (2021)
	Giraff Robot	Joose et al. (2021)

▪ *Conclusiones del análisis de los estudios experimentales*

A través de la SLR, se han identificado un total de 15 estudios experimentales que evalúan la interacción entre personas y robots. Estos estudios experimentales reflejan un creciente interés en el campo, con más del 50% de ellos publicados en los tres últimos años (2019-2021), lo que indica un impulso significativo en esta área de investigación y una creciente atención por parte de la comunidad científica, como se evidencia en la alta influencia de las revistas en las que se publicaron.

En términos generales, los experimentos se dividen en tres fases distintas: la fase previa a la ejecución de la tarea, la fase durante la ejecución y la fase posterior a la ejecución. Sin embargo, es importante destacar que no se ha identificado un modelo de evaluación validado comúnmente aceptado, ya que cada uno de estos estudios experimentales utiliza un enfoque y proceso de evaluación diferente. Con el fin de facilitar la comparación y el contraste entre ellos, se ha elaborado la Tabla 20.

Un hallazgo relevante es que ninguno de los estudios emplea una evaluación experta utilizando herramientas como heurísticas. Además, en la fase previa a la ejecución de tareas, la mayoría de los estudios no recopila datos, ya sea desde la perspectiva de personas expertas o de personas usuarias. Esto es esencial, dado que una experiencia de interacción se compone de emociones que se experimentan antes, durante y después de la interacción.

En la fase de ejecución, se observa una variedad de robots utilizados, principalmente brazos robóticos según se define en la Tabla 19. No obstante, hay casos, como el estudio de

Baskaran et al. (2019), donde se realiza la evaluación utilizando el *software Siemens Process*, prescindiendo completamente de un robot físico. Similarmente, otros estudios, como el de Colim et al. (2021), se centran en *workstations*, mientras que Almeida et al. (2020) se enfocan en las interfaces sin mencionar el uso de un robot. El desempeño se mide en siete de las experimentaciones, mientras que solo una de ellas emplea medidas fisiológicas. Siete estudios también incorporan observaciones durante la tarea para recopilar información cualitativa sobre la interacción.

En la fase posterior a la ejecución, la percepción de las personas participantes se recoge principalmente a través de cuestionarios, utilizados en trece ocasiones. En un caso, se realiza una entrevista adicional para obtener una comprensión más profunda de la percepción de la persona usuaria.

Este análisis de los estudios experimentales subraya la diversidad de enfoques y métodos utilizados en la evaluación de la HRI, lo que indica la necesidad de desarrollar un marco de evaluación más estandarizado y validado en este campo en crecimiento.

Tabla 20: Visión general y comparativa de los estudios experimentales identificados mediante la SLR.

Ref.	PRE						DURANTE								POST			
	Evaluación experta		Percepción de las personas participantes				Ejecución de la tarea								Percepción de las personas participantes			
							Cuestionario		Entrevista		Uso de robot		Quantitative objective measures					
			Desempeño		Medidas fisiológicas								Cuestionario		Entrevista			
Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No	
Schillaci et al. (2013)																		
Daniel et al. (2013)																		
Harriott et al (2013)																		
Lasota & Shah (2015)																		
Danielsson et al. (2017)																		
Aromaa et al. (2018)																		
Tang & Webb (2018)																		
Baskaran et al. (2019)																		
Kildal et al. (2019)																		
Almeida et al. (2020)																		
Hietanen et al. (2020)																		
Pantano et al. (2020)																		
Beschi et al. (2020)																		
Colim et al. (2021)																		
Joosse et al. (2021)																		

2.2.4 Resultados a la pregunta de investigación 2: ¿Qué factores humanos incluye y cómo los evalúa?

Los factores humanos son la disciplina científica que se ocupa de la interacción entre las personas y los artefactos, así como del diseño de sistemas en los que las personas participan (Helander, 1997). El propósito es adaptar sistemas, trabajos, productos y entornos a las capacidades y limitaciones físicas y mentales de las personas (Helander, 1997). Según Beith (1999) los factores humanos se centran en la usabilidad del sistema y en el diseño de interfaces del sistema para optimizar la capacidad de las personas usuarias para realizar sus tareas sin errores en un tiempo razonable y, por lo tanto, aceptar el sistema como una herramienta útil. Considerar la aplicación de los principios de los factores humanos conduce a diseños que son más seguros, más aceptables, más cómodos y más efectivos para llevar a cabo las tareas asignadas (Beith, 1999).

Esta sección aborda la segunda pregunta de investigación: ¿Qué factores humanos están involucrados en los estudios experimentales y cómo se evalúan? Para responder a esta pregunta, se ha recopilado en la Tabla 21 un análisis exhaustivo de los factores humanos considerados en estos estudios, junto con los métodos empleados para su evaluación. Se han observado por un lado medidas de desempeño. También se ha observado un enfoque considerable en el análisis de la postura de los individuos. Para ello, se han identificado seis técnicas específicas que incluyen la evaluación de la carga postural, la varianza postural, el movimiento total, la magnitud vectorial, así como los métodos RULA y lesión por esfuerzo repetitivo (RSI, por sus siglas en inglés). Por otro lado, en lo que respecta a los factores asociados al robot, se han detectado distintos aspectos a considerar, entre ellos: el antropomorfismo, la animación, la simpatía, la percepción de inteligencia, la percepción de seguridad, la usabilidad y la capacidad de aprendizaje. La medición de estos factores ha involucrado la utilización de dos cuestionarios principales, el cuestionario Godspeed y el cuestionario SUS. Finalmente, se han identificado factores intrínsecos a la persona, los cuales han sido igualmente evaluados en estos estudios. Entre estos factores se encuentran la confianza, la satisfacción, la carga mental, el estrés físico y mental, así como el riesgo percibido. Estos resultados ofrecen una comprensión más profunda de la variedad de factores humanos considerados y los métodos empleados para su evaluación en el contexto de la colaboración entre personas y robots.

2.2.4.1 Desempeño

En primer lugar, se destaca que el desempeño ha sido uno de los factores más evaluados; se refiere a cómo las personas realizan su tarea. El desempeño de las personas usuarias está moldeado por sus capacidades (por ejemplo, memoria, atención, flexibilidad), y es la consecuencia de los factores humanos presentes en el sistema. Por lo tanto, un enfoque centrado en el ser humano implica tener en cuenta también consideraciones de desempeño. Los indicadores en este grupo son aquellos que se reflejan directa o indirectamente en el desempeño humano y, por lo tanto, brindan información sobre los factores humanos y la UX. Se han identificado ocho indicadores para evaluarlo en siete de los quince estudios.

- i) **El tiempo de ejecución de la tarea** se evalúa en siete de los quince estudios (47%), lo que lo convierte en el factor más ampliamente evaluado entre los estudios identificados. Además, todos los estudios que abordan el desempeño incluyen esta medida en sus análisis.
- ii) **El número de interacciones realizadas** se registra en un estudio, específicamente en el estudio de Daniel et al. (2013). Esta medida refleja la calidad de la UI y proporciona información sobre las posibilidades de ingreso incorrecto de datos (Daniel et al., 2013).

- iii) **Los errores** se cuantifican en una ocasión, en el estudio de Almeida et al. (2020).
- iv) **El tiempo de la inactividad del robot** se registra en dos estudios, concretamente en las investigaciones de Lasota & Shah, (2015) y Hietanen et al. (2020).
- v) **El tiempo de inactividad de la persona** se evalúa en un estudio, específicamente en la investigación de Lasota & Shah (2015).
- vi) **La variabilidad en los tiempos de producción** se evalúa en un estudio, concretamente en la investigación de Colim et al. (2021).
- vii) **El ratio de producción** se evalúa en un estudio, específicamente en la investigación de Colim et al. (2021). Este indicador clave mide el desempeño en función del número de piezas producidas por hora durante un intervalo de tiempo específico (Colim et al., 2021).
- viii) **La comparación entre el tiempo requerido para concluir la tarea con y sin la intervención del robot** se evalúa en un único estudio, concretamente en la investigación de Beschi et al. (2020). Este indicador busca determinar si la productividad humana también se ve influenciada por el movimiento del robot durante tareas que no están sincronizadas (Beschi et al., 2020).

2.2.4.2 Factores relacionados con el robot

Los factores relacionados con el robot abarcan las características que describen la naturaleza del sistema y, por lo tanto, influyen en la percepción humana. Los factores relacionados con el robot son evaluados en cinco de los quince estudios. En total, se han identificado siete factores:

- i) **El antropomorfismo** se evalúa mediante la aplicación del cuestionario Godspeed desarrollado por Bartneck et al. (2008) en dos estudios. Como afirman Bartneck et al. (2008), el antropomorfismo denota la tendencia a atribuir formas, atributos o conductas propias de la naturaleza humana a entidades no humanas, como es el caso de los robots.
- ii) **La animación** es objeto de evaluación en los mismos dos estudios a través del cuestionario Godspeed, diseñado por Bartneck et al. (2008). En consonancia con las observaciones de Bartneck et al. (2008), un propósito recurrente en la comunidad investigadora en robótica radica en conferir a los robots un grado de semejanza con las formas vivas.
- iii) **La simpatía** es asimismo objeto de medición en los mismos dos análisis, empleando el cuestionario Godspeed. Como reflexión compartida por Bartneck et al. (2008), la configuración de impresiones propicias hacia los individuos se encuentra influida en parte por los modos de comportamiento visual y vocal que estos exhiben. Conforme a la misma fuente, las primeras impresiones favorables hacia una entidad suelen conllevar a apreciaciones ulteriores más benévolas en relación con dicha entidad.
- iv) **La inteligencia percibida** es igualmente evaluada mediante el cuestionario Godspeed. Los robots interactivos enfrentan enormes desafíos para actuar de manera inteligente. Las razones pueden rastrearse hasta el campo de la inteligencia artificial (IA). Los comportamientos de los robots se basan en métodos y conocimientos desarrollados con IA (Bartneck et al., 2008).
- v) **La seguridad percibida** es objeto de evaluación en tres investigaciones. En dos de ellas, esta dimensión es nuevamente medida mediante el cuestionario Godspeed. Un aspecto central en la interacción de los robots con personas es su seguridad, como señalan Bartneck et al. (2008). Este tema ha sido ampliamente abordado en la literatura sobre robótica, especialmente en lo

que respecta a los sistemas y estándares establecidos tanto para robots industriales como para aquellos diseñados para brindar servicios en entornos domésticos. En el estudio realizado por Lasota & Shah (2015) se empleó un cuestionario auto-generado de cuatro ítems para medir la percepción de seguridad.

- vi) **La usabilidad** es evaluada en dos instancias distintas. En el estudio realizado por Danielsson et al. (2017), se empleó el cuestionario System Usability Scale (SUS), diseñado por Brooke (1996). En contraste, en la investigación llevada a cabo por Almeida et al. (2020) se utilizó un cuestionario basado en el Cuestionario de Satisfacción de Usabilidad de Computadoras IBM.
- vii) **La aprendibilidad** se evalúa en una ocasión, concretamente en el estudio de Danielsson et al. (2017), donde se hace uso del cuestionario SUS. Según Joyce (2019), la noción de aprendibilidad contempla la facilidad con la que las personas usuarias pueden realizar una tarea en su primer encuentro con la interfaz, así como la cantidad de repeticiones necesarias para alcanzar eficiencia en dicha tarea.

En resumen, los siete factores relacionados con la interacción robot-humano han sido evaluados a través de la implementación de cuatro cuestionarios distintos, detallados de la siguiente manera:

- i) **Cuestionario Godspeed** (Bartneck et al., 2008) - empleado con el propósito de medir las facetas de antropomorfismo, animación, simpatía, inteligencia percibida y seguridad percibida. Este cuestionario fue utilizado en dos ocasiones específicas, es decir, en los estudios realizados por Schillaci et al. (2013) y Joosse et al. (2021).
- ii) **Cuestionario SUS** (Brooke, 1996) - para medir usabilidad y simpatía. Se utilizó una vez, es decir, en el estudio de Danielsson et al. (2017).
- iii) **Un cuestionario basado en el Cuestionario de Satisfacción de Usabilidad de Computadoras IBM**, se utilizó una vez, en el estudio de Almeida et al. (2020), para medir usabilidad y satisfacción.
- iv) Otro cuestionario autogenerado también fue utilizado en el estudio de Lasota & Shah (2015) para medir la satisfacción con los robots como compañeros de equipo y para determinar la seguridad y comodidad percibidas.

2.2.4.3 Factores relacionados con la persona

Esto abarca las respuestas emocionales resultantes de la HRI, las cuales evalúan la calidad hedónica del sistema (Diefenbach et al., 2014). Ocho de los quince estudios consideraron factores relacionados con las emociones. En total, se han identificado cinco factores:

- i) **La confianza** se mide una vez en el estudio llevado a cabo por (Daniel et al., 2013). Los autores plantearon a las personas participantes algunas preguntas adaptadas de las Iniciativas de Accesibilidad Web (WAI) para la Evaluación de la Usabilidad del Sitio (Shawn Lawton, 2003).
- ii) **La satisfacción** se ha evaluado en dos estudios, concretamente, en los de Lasota & Shah (2015) y Almeida et al. (2020). En el primero, los autores emplearon un cuestionario que evaluaba la satisfacción con los robots como compañeros de equipo; en el segundo, se utilizó un cuestionario basado en el Cuestionario de Satisfacción de Usabilidad de Computadoras IBM.
- iii) La evaluación de **la carga mental de trabajo** se lleva a cabo en tres estudios: los realizados por Aromaa et al. (2018), Pantano et al. (2020) y Harriott et al. (2013) En los tres casos, se empleó el cuestionario NASA-Task Load Index

(NASA-TLX) (Hart & Staveland, 1988). Además, el estudio de Harriott et al. (2013) enriqueció su enfoque incorporando medidas fisiológicas, tales como la frecuencia cardíaca y la variabilidad de la misma.

- iv) El estrés físico y mental se mide una vez, en el estudio realizado por Hietanen et al. (2020), utilizando un cuestionario generado por los propios investigadores.
- v) El riesgo percibido también es evaluado una vez, en el estudio realizado por Beschi et al. (2020), empleando un cuestionario generado por los propios investigadores.

En cuanto a la postura, relacionada con la antropometría y la biomecánica, esta se enfoca en eliminar prácticas laborales perjudiciales e inseguras, y tiene como objetivo estudiar las capacidades y limitaciones humanas con el fin de adaptar la tarea a la persona mientras se minimiza la fatiga (Cardoso et al., 2021). Cuatro de los quince estudios analizan la postura de la persona. Se han identificado seis indicadores a partir de estos estudios:

- i) **La carga postural** se mide en un estudio, concretamente en el estudio de Harriott et al. (2013). Se refiere al porcentaje de tiempo que las personas participantes pasaron con la flexión de sus troncos a un ángulo de más de 45º respecto a la vertical (Harriott et al., 2013). Cuanto más tiempo pase un participante con flexión severa del tronco, mayor será la carga física (Kuijer et al., 1999).
- ii) **La variabilidad en la postura** se mide en un único estudio, en el de Harriott et al. (2013).
- iii) **El movimiento total** se mide en un estudio, en el de Harriott et al. (2013). Se presenta como el número total de veces que el participante se puso de pie y se agachó.
- iv) **La magnitud vectorial** se mide en el estudio de Harriott et al. (2013). Se trata de una medida de la actividad física general y combina la aceleración de los tres ejes de movimiento. La magnitud vectorial mide el movimiento físico de las personas participantes en el área de evaluación.
- v) **La Evaluación Rápida del Miembro Superior (RULA)** es un método de observación (David, 2005) para evaluar trastornos musculoesqueléticos relacionados con el trabajo físico en los miembros superiores (McAtamney & Corlett, 1993). Su aplicación implica la evaluación de la postura de la persona trabajadora, así como las fuerzas ejercidas, la repetitividad de los movimientos y las cargas externas (por ejemplo, manejo de materiales pesados). Se mide en tres estudios: Aromaa et al. (2018), Tang & Webb (2018) y Colim et al. (2021).
- vi) **El Índice de Tensión Revisado (RSI)** también se mide en el estudio de Colim et al. (2021). El RSI consiste en un modelo de cinco variables con multiplicadores continuos. Las cinco variables/factores de riesgo medidos son la intensidad del esfuerzo (fuerza), los esfuerzos por minuto (frecuencia), la duración por esfuerzo, la postura mano-muñeca y la duración de la tarea por día (Garg et al., 2017).

Tabla 21: Factores humanos evaluados y las herramientas utilizadas.

Factor	Técnica	Referencia	
Desempeño	Tiempo de ejecución de la tarea	Daniel et al. (2013), Harriott et al. (2013), Lasota & Shah (2015), Almeida et al. (2020), Beschi et al. (2020), Hietanen et al. (2020), Colim et al. (2021)	
	Número de interacciones realizadas	Daniel et al. (2013)	
	Errores	(Almeida et al., 2020)	
	Tiempo de inactividad del robot	Lasota & Shah (2015), Hietanen et al. (2020)	
	Tiempo de inactividad de la persona	Lasota & Shah (2015)	
	Variabilidad en los tiempos de producción	Colim et al. (2021)	
	Índice de producción	Colim et al. (2021)	
	Cociente entre el tiempo necesario para completar la tarea con y sin el robot	Beschi et al. (2020)	
Factores relacionados con el robot	Antropomorfismo	Cuestionario Godspeed Schillaci et al. (2013), Joosse et al. (2021)	
	Animación	Cuestionario Godspeed Schillaci et al. (2013), Joosse et al. (2021).	
	Simpatía	Cuestionario Godspeed Schillaci et al. (2013), Joosse et al. (2021).	
	Inteligencia percibida	Cuestionario Godspeed Schillaci et al. (2013), Joosse et al. (2021).	
	Seguridad percibida	Cuestionario Godspeed	Schillaci et al. (2013), Joosse et al. (2021).
		Cuestionario	Lasota & Shah (2015)
	Usabilidad	Cuestionario SUS	Danielsson et al. (2017)
		Cuestionario basado en IBM Computer Usability Satisfaction Questionnaire	Almeida et al. (2020)
Aprendibilidad	Cuestionario SUS	Danielsson et al (2017)	
Factores relacionados con la persona	Confianza	Cuestionario sobre la confianza	Daniel et al. (2013)
		Cuestionario	Lasota & Shah (2015)
	Satisfacción	Cuestionario basado en IBM Computer Usability Satisfaction Questionnaire	Almeida et al. (2020)
		Carga mental	NASA-TLX
	Ritmo cardíaco		Harriott et al. (2013)
	Variabilidad del ritmo cardíaco		Harriott et al. (2013)
	Estrés físico y mental	Cuestionario sobre estrés físico y mental	Hietanen et al. (2020)
	Riesgo percibido	Cuestionario sobre el riesgo percibido	Beschi et al. (2020)
	Postura	Carga postural	Harriott et al. (2013)
		Varianza postural	Harriott et al. (2013)
		Movimiento total	Harriott et al. (2013)
		Magnitud vectorial	Harriott et al. (2013)
		RULA	Aromaa et al. (2018), Tang & Webb (2018), Colim et al. (2021)
		RSI	Colim et al (2021)

2.2.4.4 Clasificación según las técnicas y herramientas utilizadas para la evaluación

Se han contabilizado en la Tabla 22 las herramientas utilizadas en los estudios experimentales identificados y se han clasificado según si se tratan de medidas objetivas o subjetivas, y si son cuantitativas o cualitativas.

Tabla 22: Clasificación de las herramientas identificadas.

		Herramienta	Referencia
Medidas objetivas	Medidas cuantitativas	Tiempo de ejecución de la tarea	Daniel et al. (2013), Harriott et al. (2013), Lasota & Shah (2015), Almeida et al. (2020), Beschi et al. (2020), Hietanen et al. (2020), Colim et al. (2021).
		Número de interacciones realizadas	Daniel et al. (2013)
		Errores	Almeida et al. (2020)
		Precisión de seguimiento de la trayectoria	Almeida et al. (2020)
		Relación entre las interacciones con la pantalla táctil y el teclado	Daniel et al. (2013)
		Movimiento simultáneo	Lasota & Shah (2015)
		Distancia media de separación	Lasota & Shah (2015), Beschi et al. (2020)
		Tiempo de inactividad del robot	Lasota & Shah (2015), Hietanen et al. (2020)
		Tiempo de inactividad de la persona	Lasota & Shah (2015)
		Variabilidad en los tiempos de producción	Colim et al. (2021)
		Índice de producción	Colim et al. (2021)
		Cociente entre el tiempo necesario para completar la tarea con y sin el robot	Beschi et al. (2020)
		Carga postural	Harriott et al. (2013)
		Varianza postural	Harriott et al. (2013)
		Movimiento total	Harriott et al. (2013)
	Magnitud vectorial	Harriott et al. (2013)	
	Medidas de postura	RULA	Aromaa et al. (2018), Tang & Webb (2018), Colim et al. (2021).
		RSI	Colim et al. (2021)
		Monitor torácico de electrocardiografía BioHarness	Harriott et al. (2013)
		Siemens process	Baskaran et al. (2019)
Medidas subjetivas	Medidas cuantitativas	Cuestionario Godspeed	Schillaci et al. (2013), Joosse et al. (2021).
		Cuestionario Trust in Automation	Daniel et al. (2013)
		Cuestionario (Satisfacción con el robot como compañero; comodidad y seguridad percibida)	Lasota & Shah (2015)
		Cuestionario SUS	Danielsson et al. (2017), Tang & Webb (2018)
		Cuestionario de interés general	Danielsson et al. (2017)

	NASA-TLX	Harriott et al. (2013), Aromaa et al. (2018), Pantano et al. (2020).
	Cuestionario basado en IBM Computer Usability Satisfaction Questionnaire y preguntas de presencia.	Almeida et al. (2020)
	Cuestionario sobre estrés físico y mental	Hietanen et al. (2020)
	Cuestionario sobre el riesgo percibido	Beschi et al. (2020)
	Cuestionario sobre cambios en la configuración de la planificación y con qué se sienten más cómodos	Beschi et al. (2020)
	Cuestionario sobre la percepción de la persona trabajadora (impacto de la robótica en el contexto laboral, esfuerzo percibido asociado a las tareas y evaluación global del puesto de trabajo)	Colim et al. (2021)
Medidas cualitativas	Observación	Danielsson et al. (2017), Kildal et al. (2019)
	Calificación verbal sobre la carga de trabajo (auditiva, visual, del habla, táctil y motor)	Harriott et al. (2013)

2.3 Últimos avances en la literatura

En esta sección se han incorporado los avances más recientes en la literatura. En este contexto, se han identificado once trabajos adicionales siguiendo la misma lógica utilizada en la SLR. De estos once trabajos, cuatro se clasifican como marcos teóricos, otros cuatro como herramientas, y tres como estudios experimentales.

2.3.1 Marcos teóricos

A continuación, se presentan y analizan los cuatro marcos teóricos, que son:

- i) *Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review* (Gaultieri et al., 2021)
- ii) *Evaluating quality in human-robot interaction: A systematic search and classification of performance and human-centered factors, measures and metrics towards an industry 5.0* (Coronado et al., 2022)
- iii) *Designing human-robot collaboration (HRC) workspaces in industrial settings: A systematic literature review* (Simões et al., 2022)
- iv) *Human factors in cobot era: a review of modern production systems features* (Faccio et al., 2022)

2.3.1.1 *Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review (Gaultieri et al., 2021)*

En el estudio realizado por Gaultieri et al. (2021) se llevó a cabo una SLR con el objetivo de analizar el estado actual de la seguridad y la ergonomía en la robótica colaborativa industrial. Para organizar los temas emergentes identificados, se utilizaron clústeres y sub-clústeres (Figura 21) basados en una división primaria en "Seguridad" y "Ergonomía".

En la categoría de seguridad, se identificaron dos clústeres principales: "Evitar el contacto" y "Detección y mitigación del contacto". Estos se enfocan en salvaguardar la

seguridad de las personas operarias al prevenir y controlar contactos peligrosos entre personas y robots. Por otro lado, en la categoría de ergonomía, se diferenciaron entre "Ergonomía física" y "Ergonomía cognitiva y organizacional", que abordan aspectos relacionados con la reducción de la carga biomecánica y el bienestar mental de las personas operarias, así como la optimización de los sistemas organizativos y políticos.

En los subclústeres de nivel inferior se identificaron los principales argumentos que se asignaron a cada uno de los cuatro clústeres. El clúster 1, denominado "Evitar el contacto", se enfoca en salvaguardar la seguridad de las personas operarias al prevenir contactos peligrosos mediante metodologías y sistemas preventivos. Dentro de este clúster identificaron varios subclústeres que abordan aspectos específicos de la seguridad en la interacción entre personas y robots. El sub-clúster de "Planificación y Control del Movimiento" se centra en desarrollar métodos para predecir la ocupación del espacio humano, modificar trayectorias de forma adaptativa y controlar el movimiento de los robots. El sub-clúster de "Sistemas de Sensores para el Seguimiento de Objetos" se enfoca en el desarrollo e integración de sistemas de monitoreo y visión por computadora para localizar y controlar a las personas en el entorno de trabajo, así como reconocer gestos. El sub-clúster de "Gestión de la Seguridad" se dedica a la creación de procedimientos, metodologías, métricas e índices para la prevención de contactos peligrosos. Además, realizaron contribuciones menores en los sub-clústeres de "Simulación y Modelado", "Inteligencia Artificial", "Sistemas de Asistencia" y "Estudios de Casos y Aplicaciones". Estos avances han abordado temáticas clave para mejorar la seguridad en las interacciones entre personas y robots.

El clúster 2, denominado "Detección y Mitigación de Contactos", tiene como objetivo garantizar la seguridad de las personas operarias en términos de riesgo mecánico mediante la reducción de la energía de colisión que puede intercambiarse durante contactos no deseados o inesperados entre personas y robots. Dentro de este clúster se han identificado varios subclústeres que abordan aspectos específicos de la detección y mitigación de contactos. El sub-clúster de "Planificación y Control del Movimiento" se centra en desarrollar estrategias de control para evitar colisiones. El sub-clúster de "Diseño de Sistemas de Robots" se dedica al desarrollo de *hardware* de robots y metodologías de diseño. El sub-clúster de "Sistemas de Sensores para la Gestión de Contactos" se enfoca en el desarrollo de dispositivos de sensores y metodologías de detección. Además, se han realizado contribuciones menores en los subclústeres de "Simulación y Modelado", "Gestión de la Seguridad" y "Estudios de Casos y Aplicaciones". Estos avances han abordado temáticas clave para mejorar la detección y mitigación de contactos en las interacciones entre personas y robots.

El clúster 3, denominado "Ergonomía Física", se centra en la aplicación de principios ergonómicos para reducir la carga biomecánica de las actividades en la HRI en entornos industriales. Se utilizan robots colaborativos como herramientas avanzadas para mejorar el bienestar físico de las personas operarias. Dentro de este clúster, se han identificado varios subclústeres que abordan aspectos específicos de la ergonomía física. El sub-clúster de "Planificación y Control del Movimiento" se enfoca en el desarrollo de estrategias de control del movimiento que consideren los factores de ergonomía física. El sub-clúster de "Estudios de Casos y Aplicaciones" se dedica al desarrollo de celdas colaborativas que mejoran la ergonomía física de las personas operarias. El sub-clúster de "Sistemas de Asistencia" se centra en el desarrollo de sistemas de interacción multimodal y tecnologías habilitadoras para mejorar las condiciones de ergonomía física. Además, se han realizado contribuciones menores en los subclústeres de "Gestión de la Seguridad", "Métricas y Pruebas" y "Diseño de Productos y Procesos". Estos avances han abordado temáticas clave para mejorar la ergonomía física, como la planificación de tareas, el diseño de productos y procesos, y la gestión de la seguridad en la HRI.

El clúster 4, denominado "Ergonomía Cognitiva y Organizacional", aborda dos aspectos importantes en la HRI en entornos industriales. La ergonomía cognitiva se centra en minimizar el estrés mental y el malestar psicológico que puedan experimentar las personas operarias al compartir el espacio de trabajo con robots. Por otro lado, la ergonomía organizacional se refiere a la optimización de los sistemas técnicos y sociales en términos de estructuras organizativas, políticas y procesos. Dentro de este clúster, se han identificado varios subclústeres que abordan aspectos específicos de la ergonomía cognitiva y organizacional. El subclúster de "Planificación y Control del Movimiento" se enfoca en el desarrollo de arquitecturas de control que mejoren la aceptabilidad del sistema desde el punto de vista cognitivo. El subclúster de "Simulación y Modelado" se dedica al desarrollo de entornos virtuales para evaluar la aceptabilidad cognitiva de la HRI. El subclúster de "Sistemas de Asistencia" se centra en el desarrollo de sistemas de apoyo cognitivo, como soportes audiovisuales y modelos de enseñanza-aprendizaje-colaboración basados en lenguaje natural. Además, se han realizado contribuciones en los subclústeres de "Métricas y Pruebas" y "Estrategias de Programación de Tareas", que abordan la evaluación de la aceptabilidad del robot y la asignación de tareas basada en la confianza. Estos avances han abordado temáticas clave para mejorar la ergonomía cognitiva y organizacional, como la evaluación de la aceptabilidad, el control basado en aspectos cognitivos y la simulación de entornos virtuales.

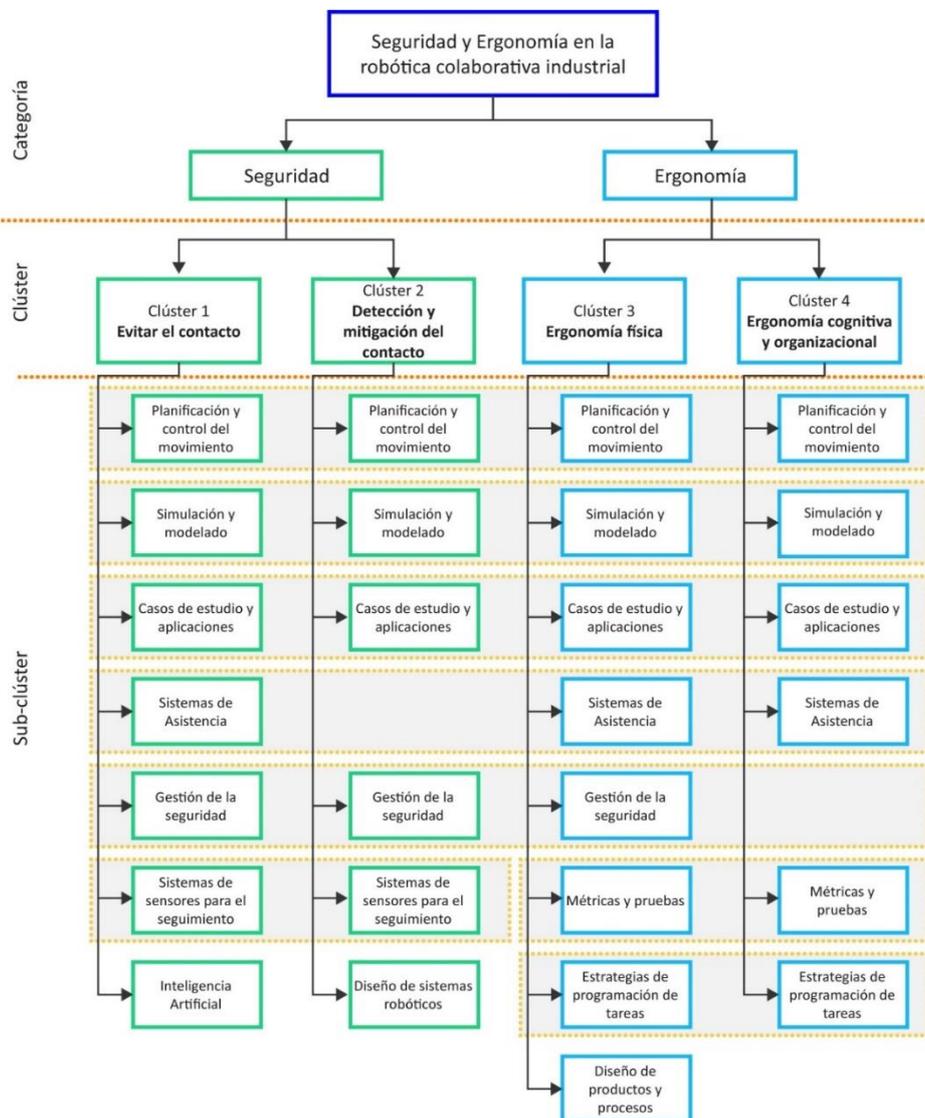


Figura 21: Identificación de categorías, clústeres y subclústeres. Adaptado de Gualtieri et al. (2021).

Siendo así, los resultados de Gualtieri et al. (2021) muestran que la categoría de investigación más desarrollada es la seguridad, aunque la ergonomía ha crecido significativamente en los últimos años (especialmente la parte cognitiva y organizativa). En particular, los resultados muestran que la mayoría de los temas de mayor importancia están más relacionados con los aspectos de seguridad que con la ergonomía. En lo que respecta a los trabajos de seguridad, hay un mayor interés en la prevención (que estudia las medidas de seguridad que operan antes de que ocurran posibles contactos entre personas y robots) que en la protección (que estudia las medidas de seguridad que buscan minimizar los efectos de los contactos persona-robot después de que ocurra la interacción física). La evitación de colisiones es el clúster más grande caracterizado por una tasa de producción de documentos lenta pero constante. Sin embargo, aunque la ergonomía presentó la mayor producción de documentos en los últimos años, los temas relacionados se consideran emergentes, pero aún no maduros (Gualtieri et al., 2021).

2.3.1.2 Evaluating quality in human-robot interaction: A systematic search and classification of performance and human-centered factors, measures and metrics towards an industry 5.0 (Coronado et al., 2022)

El estudio de Coronado et al. (2022) presenta una revisión bibliográfica de los diferentes aspectos relativos a la medición de la calidad en HRI en entornos industriales. Presentan una visión general de los factores, métricas y medidas utilizadas en la comunidad robótica para evaluar el rendimiento y los aspectos de calidad del bienestar de las personas operarias. Presentan una taxonomía de aspectos de desempeño y un diagrama de Venn de los factores humanos más comunes en HRI.

En la Tabla 23 se identifican las métricas identificadas por Coronado et al. (2022), que se categoriza según:

- Las medidas de **comportamiento temporal** indican los tiempos de respuesta y procesamiento que una persona, un robot o una combinación de personas y robots necesitan para realizar sus funciones, una subtarea o una tarea completa. Ejemplos de estas métricas son el tiempo de inactividad humano, el tiempo de procesamiento del algoritmo, el tiempo de colaboración y el tiempo de finalización de la tarea.
- Las **medidas de proceso** son una agregación de hechos generados desde el inicio hasta el final de una tarea o subtarea, así como elementos relacionados con los costes, el diseño del espacio de trabajo, la seguridad o la calidad del producto. Ejemplos de estas métricas son el número de errores y el número de ensamblajes alcanzados.
- Las **medidas fisiológicas** son valores obtenidos a partir de medidas corporales que ayudan a comprender el estado actual del ser humano (por ejemplo, la aceleración de las articulaciones humanas y la frecuencia cardíaca).
- Las **medidas físicas humano-robot** son valores obtenidos a partir de sensores que indican el estado actual de la interacción (por ejemplo, la distancia entre la persona y el robot).
- **Métricas de eficiencia.** Según la norma ISO 9241, la eficiencia es la "relación entre los recursos (inputs) utilizados y los resultados (outputs) obtenidos". Coronado et al. (2022) definen las métricas que evalúan la eficiencia como relaciones de entrada/salida. La idea principal de las métricas de eficiencia es evaluar si los sistemas de HRI están "haciendo las cosas bien". Por tanto, estas métricas evalúan el progreso hacia la consecución de los objetivos definidos. En consecuencia, la pregunta típica a la que intentan responder es cómo de bien se utilizan los recursos (tiempo, costes, materiales).

- Las **métricas de efectividad** expresan la relación entre los resultados reales u obtenidos y los resultados programados, deseados o previstos. La idea principal de las métricas de efectividad es evaluar si los sistemas HRI están "haciendo las cosas bien". Por lo tanto, estas métricas evalúan la precisión y la exhaustividad con la que los sistemas de HRI alcanzan los objetivos especificados.

Tabla 23: Categorización orientada al rendimiento para las métricas obtenidas por Coronado et al. (2022).

Categoría	Métrica	Nivel de aplicación		
		Persona	Robot	Persona y robot
Comportamiento temporal	Tiempo de procesamiento del algoritmo			
	Tiempo de ensamblado	■	■	
	Tiempo promedio para ejecutar la tarea	■	■	
	Tiempo de colaboración			
	Tiempo de cooperación			
	Tiempo de coordinación			
	Duración (tecnología de comunicación)			
	Tiempo libre			
	Retrasos funcionales			
	Tiempo de acción humana	■		
	Tiempo de inactividad humana	■		
	Tiempo de operación humana	■		
	Tiempo de inactividad			
	Esfuerzo de interacción	■		
	Tiempo de interacción			
	Tiempo desatendido		■	
	Tiempo de reacción	■	■	
	Tiempo de reprogramación			
	Tiempo de respuesta	■	■	
	Tiempo de acción del robot		■	
	Demanda de atención del robot		■	
	Retardo funcional del robot		■	
	Tiempo de inactividad del robot		■	
	Tiempo de funcionamiento del robot		■	
	Tiempo de preparación	■	■	
	Latencia del sistema			
Tiempo de ejecución de la tarea	■	■		
Tiempo total de ensamblaje				
Tiempo de funcionamiento total				
Tiempo de procesamiento	■	■		
Medidas del proceso	Coste de la cadena de montaje		■	
	Coste del sistema HRC		■	
	Desviación de la trayectoria definida		■	
	Número de trabajadores cualificados en la línea			
	Seguridad basada en el número de colisiones			
Medidas fisiológicas	Recuentos de asignación de tareas			
	Aceleración de las articulaciones humanas			
	Bioseñales (temperatura, tacto, etc.)			
	Carga biomecánica	■		
	Mejora de la ergonomía	■		
Actividad muscular	■			
Fatiga muscular del brazo	■			

Categoría	Métrica	Nivel de aplicación		
		Persona	Robot	Persona y robot
Medidas físicas de HR	Manipulabilidad muscular	■		■
	Comportamiento ocular	■		
	Respuesta potencial de la piel	■		
	Conductancia cutánea	■		
	Promedio por minuto de la distancia entre la mano humana y la mano del robot		■	
	Dirección de la reacción		■	
	Riesgo relacionado con la fuerza basado en HIC		■	
	Distancia persona-robot		■	
	Torques articulares de sobrecarga humana para todo el cuerpo	■		
	Disponibilidad	■	■	
Eficiencia	Velocidad media del robot		■	
	Concentración o atención sostenida	■		
	Actividad simultánea			
	Movimiento simultáneo			
	Duración del ciclo	■	■	
	Grado de colaboración			
	Eficiencia económica		■	
	Índice de evaluación económica	■	■	
	Eficiencia basada en la velocidad media de y efector			
	Eficiencia basada en el tiempo de movimiento neto			
	Varianza de la carga energética entre las personas trabajadoras		■	
	Grado de utilización (tecnología de la comunicación)			
	Eficiencia del trabajo en equipo con interfaz			
	Eficiencia del diseño			
	Velocidad media del efector final			
Efectividad	Tiempo total de movimiento	■	■	
	Producción			
	Velocidad del robot		■	
	Índice de evaluación técnica	■	■	
	Precisión			
	Error medio de predicción		■	
	Tasa de interacciones falsas negativas			
	Tasa de interacciones falsas positivas			
	Precisión de la interacción			
	Nivel de asignación			
Nivel de interacción				
Efectividad global de los equipos	■			
Efectividad global de los equipos para HRI				
Error de predicción		■		
Índice de evaluación cualitativa	■	■		
Fallo humano en tiempo real	■			
Fallo del robot en tiempo real		■		

2.3.1.3 *Designing human-robot collaboration (HRC) workspaces in industrial settings: A systematic literature review (Simões et al., 2022)*

El estudio de Simões et al. (2022) proporciona una SLR sobre el diseño de espacios de HRC en entornos industriales. Los autores desarrollaron un marco conceptual que representa los niveles de complejidad de los factores influyentes presentes en los contextos de HRI para el análisis de contenido. Basándose en este marco conceptual, Simões et al. (2022) presentan una serie de pautas y recomendaciones de los artículos analizados en tres categorías:

- Categoría 1: El primer nivel de complejidad se centra en un único factor influyente específico en la HRI. Esta categoría se divide en dos subcategorías: la persona operaria y la tecnología.
- Categoría 2: El segundo nivel de complejidad aborda recomendaciones y pautas relacionadas con el rendimiento del equipo formado por personas y robots. En esta categoría, múltiples factores influyentes están presentes en la HRI.
- Categoría 3: El tercer nivel de complejidad se enfoca en enfoques más complejos y holísticos para el diseño de sistemas de HRC. Esta categoría considera tanto los aspectos humanos como los tecnológicos en el diseño de sistemas HRC efectivos y seguros.

La Categoría 1 se divide en dos subcategorías que abordan diferentes aspectos de la HRC, centrándose en la interacción entre ambos. La subcategoría 1a se enfoca en el papel de la persona operaria en esta colaboración y examina los procesos cognitivos y sociales que influyen en esta interacción. Su objetivo principal es comprender cómo mejorar esta colaboración para lograr resultados eficientes y seguros.

Dentro de la subcategoría 1a, se investiga en primer lugar la confianza como aspecto clave. Se analiza cómo se desarrolla y establece la confianza entre la persona operaria y el robot, ya que esta es fundamental para una colaboración exitosa. La persona operaria necesita creer en la capacidad del robot para llevar a cabo tareas de manera precisa, confiable y segura. En este sentido, se examinan factores que influyen en la confianza, como la experiencia previa de la persona operaria con robots, la calidad de las interacciones pasadas y la confiabilidad demostrada por el robot. Otro aspecto estudiado en esta subcategoría es la atribución de culpa. Se investiga cómo la persona operaria atribuye la responsabilidad de los errores o fracasos durante la colaboración, ya sea al robot o a sí mismo. Comprender cómo se atribuye la culpa resulta crucial para mejorar la colaboración y evitar conflictos potenciales. Se exploran factores como la claridad de las responsabilidades asignadas al operador y al robot, la transparencia en las acciones del robot, y la capacidad de aprendizaje y adaptación del robot para minimizar errores. Asimismo, se examina la aceptación de la tecnología en esta subcategoría. Se analiza el grado en que la persona operaria se siente cómodo trabajando con el robot y está dispuesto a utilizarlo en su entorno laboral. Se evalúa la actitud y la disposición de la persona operaria para adoptar y utilizar nuevas tecnologías. Esto implica considerar factores como la usabilidad del robot, la familiaridad con la tecnología y los beneficios percibidos de la HRC. En cuanto a los procesos sociales, se investigan las dinámicas de relación entre la persona operaria y el robot. Se examina cómo se establecen y mantienen estas relaciones, y cómo influyen en la colaboración. Se exploran aspectos como la comunicación efectiva, la interacción fluida y la cooperación entre la persona operaria y el robot. Factores como la personalidad de la persona operaria, las preferencias de interacción y la capacidad del robot para adaptarse al estilo de trabajo de la persona operaria pueden influir en la dinámica de la relación. Además, se consideran las expectativas sociales y las normas culturales que rodean la HRC. Se analizan las percepciones y expectativas de la sociedad en general y cómo pueden influir en la forma en que se lleva a cabo y se percibe la colaboración. También se investiga cómo las normas culturales específicas de una sociedad o comunidad pueden afectar la colaboración y las preferencias en cuanto a la interacción y la comunicación.

La subcategoría 1b se enfoca en la tecnología utilizada en la HRC. Se examinan los aspectos relacionados con el diseño, desarrollo e implementación de la tecnología robótica, así como sus implicaciones en la interacción con las personas. Uno de los aspectos clave investigados en esta subcategoría es el diseño de la interfaz. Se busca desarrollar interfaces intuitivas y efectivas que permitan una comunicación fluida y eficiente entre la persona y el robot. Esto implica considerar aspectos como la visualización de información, el diseño de controles y la retroalimentación sensorial. Se exploran técnicas HCD para garantizar que la interfaz sea fácil de usar y se adapte a las necesidades y preferencias de la persona operaria. La inteligencia artificial y el aprendizaje automático son aspectos clave en esta subcategoría. Se investiga cómo utilizar algoritmos y técnicas de aprendizaje automático para mejorar las capacidades del robot, como el reconocimiento de voz, la percepción visual y la toma de decisiones. Se exploran enfoques de aprendizaje automático que permitan a los robots adaptarse a las preferencias y necesidades específicas de las personas operarias, así como a los cambios en el entorno de trabajo. La seguridad es otro aspecto crítico abordado en la subcategoría 1b. Se examinan los protocolos y las medidas de seguridad necesarias para garantizar que la interacción entre la persona y el robot sea segura y libre de riesgos. Esto implica la implementación de sistemas de detección y prevención de colisiones, la evaluación de riesgos y la adopción de estándares de seguridad adecuados. También se investigan técnicas de seguridad en la comunicación entre personas y robots, como la encriptación y la autenticación. La conectividad y la interoperabilidad son consideraciones importantes en esta subcategoría. Se busca desarrollar tecnologías que permitan la integración fluida de robots en entornos existentes y sistemas de información. Esto implica el uso de estándares de comunicación y protocolos que faciliten la interoperabilidad entre diferentes sistemas y dispositivos. Además, se exploran enfoques de conectividad robustos que permitan la comunicación confiable y en tiempo real entre la persona operaria y el robot. La evaluación y la medición del desempeño de la tecnología robótica son también temas investigados en esta subcategoría. Se desarrollan métricas y metodologías para evaluar el rendimiento de los robots en diferentes tareas y contextos de colaboración. Esto incluye el análisis de la precisión, la eficiencia y la capacidad de adaptación de los robots. Además, se investigan técnicas de retroalimentación y mejora continua para optimizar el desempeño de la tecnología robótica a lo largo del tiempo.

La Categoría 2 se enfoca en el rendimiento de los equipos conformados por personas y robots en entornos colaborativos. La investigación se centra en comprender los factores que influyen en la eficacia y eficiencia de estos equipos, así como en desarrollar estrategias y técnicas para mejorar su rendimiento. Un aspecto clave abordado en esta categoría es la coordinación entre personas y robots. Se analiza cómo los miembros del equipo interactúan y se comunican entre sí para lograr los objetivos comunes. Se examinan aspectos como la distribución de tareas y roles, la asignación de responsabilidades y la sincronización de acciones. El objetivo es encontrar enfoques de coordinación que permitan una colaboración fluida y sin conflictos, maximizando la eficiencia del equipo. La planificación y la toma de decisiones son áreas de investigación importantes en esta categoría. Se investiga cómo los equipos persona-robot pueden desarrollar estrategias y tomar decisiones efectivas en tiempo real. Se analiza cómo los robots pueden utilizar su capacidad de procesamiento y análisis de datos para apoyar la toma de decisiones humanas. Se exploran técnicas de planificación conjunta que optimicen el uso de los recursos del equipo y maximicen el logro de los objetivos. La adaptabilidad y la flexibilidad son aspectos cruciales en el rendimiento de los equipos persona-robot. Se investiga cómo los equipos pueden adaptarse a diferentes tareas, entornos y condiciones cambiantes. Esto implica el desarrollo de algoritmos y estrategias que permitan a los robots ajustar su comportamiento y respuesta según las necesidades del equipo y el contexto. Se exploran enfoques de aprendizaje automático y planificación reactiva que posibiliten una adaptación rápida y eficiente. La confianza y la percepción de los miembros del

equipo también son factores fundamentales en esta categoría. Se investiga cómo las personas y los robots pueden construir confianza mutua y comprender las habilidades y limitaciones de cada uno. Esto implica una comunicación clara de las intenciones y el estado del equipo, así como una retroalimentación efectiva. Se exploran técnicas de visualización y comunicación que faciliten la comprensión mutua y refuercen la confianza entre los miembros del equipo. La evaluación del rendimiento del equipo persona-robot es otra área de investigación importante. Se desarrollan métricas y metodologías para medir la efectividad y la eficiencia del equipo en diferentes tareas y escenarios. Esto implica analizar el desempeño individual de los miembros del equipo, así como el desempeño colectivo. Se investigan técnicas de evaluación que permitan identificar las fortalezas y debilidades del equipo y proporcionen información para su mejora continua.

La Categoría 3 se centra en enfoques integrados para el diseño de sistemas de HRC. Esta categoría aborda la necesidad de desarrollar enfoques holísticos que consideren tanto los aspectos humanos como los tecnológicos en el diseño de sistemas HRC efectivos y seguros. En primer lugar, se investiga cómo integrar de manera óptima las habilidades y capacidades de las personas y los robots para maximizar la eficiencia y la productividad del sistema. Se analizan los requisitos de las tareas y se diseñan estrategias que permitan aprovechar las ventajas de cada componente del equipo. Esto implica identificar las tareas que son más adecuadas para las personas y las tareas que pueden ser realizadas de manera más eficiente por los robots. Se buscan soluciones que permitan una colaboración fluida y complementaria entre personas y robots, optimizando la utilización de los recursos disponibles. Además, se investiga cómo diseñar interfaces de usuario intuitivas y accesibles para los sistemas HRC. Se analiza cómo presentar información de manera clara y comprensible tanto para las personas como para los robots, facilitando la comunicación y la interacción eficaz. Se exploran técnicas de visualización, retroalimentación y control que permitan a las personas usuarias comprender el estado del sistema y controlar las acciones de los robots de manera intuitiva. Se busca mejorar la experiencia de la persona usuaria y reducir la carga cognitiva asociada con la interacción con el sistema HRC. La seguridad es otro aspecto crucial en el diseño de sistemas HRC, por lo que esta categoría también aborda la integración de enfoques de seguridad en el diseño. Se investiga cómo garantizar la seguridad de las personas y los robots en entornos colaborativos. Esto implica el desarrollo de técnicas de detección y prevención de colisiones, así como la implementación de mecanismos de supervisión y control que minimicen los riesgos de lesiones o daños. Se exploran enfoques de control de seguridad que permitan una colaboración segura y confiable entre personas y robots. Otro aspecto importante es la consideración de los aspectos éticos y legales en el diseño de sistemas HRC. Se investiga cómo diseñar sistemas que respeten la privacidad, la autonomía y la dignidad de las personas, así como los principios éticos fundamentales. Se abordan cuestiones relacionadas con la responsabilidad y la rendición de cuentas en caso de accidentes o malfuncionamientos. Se exploran marcos legales y éticos que guíen el diseño y el uso de sistemas HRC de manera ética y responsable.

Siendo así, la SLR de Simões et al. (2022) demuestran que el diseño de un lugar de trabajo colaborativo implica considerar varios factores influyentes que están involucrados en la HRI. Por lo tanto, el diseño de un lugar de trabajo colaborativo debe abordarse de manera integrada y multidisciplinaria para maximizar la participación humana en la cadena de decisiones y promover el bienestar y la calidad del trabajo.

2.3.1.4 Human factors in cobot era: a review of modern production systems features (Faccio et al., 2022)

El estudio de Faccio et al. (2022) proporciona una visión general de la robótica colaborativa, considerando tres dimensiones principales: características de los robots, características de los sistemas de producción modernos y factores humanos. La interacción directa entre personas y robots puede ser tanto una ventaja como una limitación de los

sistemas colaborativos, dependiendo de cómo afecte a los factores humanos como la ergonomía y el estrés mental.

Para empezar y con un enfoque similar al de Gervasi et al. (2020) y al de Rücker et al. (2019), Faccio et al. (2022) consideran los siguientes factores humanos clave en la HRI (Tabla 24).

Tabla 24: Factores humanos en la era de los cobots. Adaptado de (Faccio et al., 2022)

Factores humanos	Descripción
Ergonomía física	Los aspectos físicos, musculoesqueléticos y sensoriales de las actividades laborales y colaborativas.
Carga mental de trabajo	Los aspectos cognitivos y emocionales relacionados con la demanda laboral (por ejemplo, complejidad de las tareas), la arquitectura cognitiva humana (por ejemplo, recursos cognitivos limitados) y las emociones y estados de ánimo humanos.
Confianza	La actitud que el robot ayuda a alcanzar un objetivo en una situación específica, relacionada con la comprensión de la persona trabajadora sobre las habilidades y limitaciones del robot (por ejemplo, seguridad percibida, confiabilidad, conciencia de la situación y transparencia, velocidad de movimiento y recogida del robot).
Aceptación	Los aspectos relacionados con las creencias subjetivas de que el robot puede mejorar el rendimiento en el trabajo y que se puede usar sin esfuerzo, también en relación con las posibles consecuencias de la introducción del cobot a nivel de la comunidad laboral (por ejemplo, cambios en el rol de las personas trabajadoras, pérdida de empleo).
Usabilidad	Los aspectos relacionados con la efectividad, eficiencia y satisfacción en la consecución de objetivos específicos.

El estudio realizado por Faccio et al. (2022) ofrece una visión amplia de cómo interactúan los factores humanos, las capacidades de los cobots y las características modernas de los sistemas de producción. Destacan que los factores humanos más prominentes son la ergonomía física y la carga mental. Estos aspectos son estudiados en relación con las capacidades de los cobots, tales como la seguridad, la inteligencia y la conectividad. Sin embargo, se observa un desequilibrio en la atención prestada a estos factores humanos en comparación con otras capacidades de los cobots, como la consistencia.

En paralelo, la investigación destaca que ciertas características de los sistemas de producción modernos, como la reconfigurabilidad y la orientación al costo, han sido ampliamente analizadas en relación con la ergonomía y la carga mental. Sin embargo, se evidencia una falta de atención hacia otros factores tanto en el ámbito de los factores humanos como en las características de los sistemas de producción.

En este sentido, Faccio et al. (2022) identifican las siguientes conclusiones:

- En relación con los Factores Humanos, la investigación ha analizado de manera equitativa diferentes aspectos, haciendo énfasis en la ergonomía física. Se reconoce que la mera observación de las directrices de seguridad no siempre es suficiente para garantizar condiciones de trabajo óptimas y, en el peor de los casos, puede resultar en condiciones de trabajo degradadas, tanto físicas como mentales. Además, la ergonomía física generalmente se considera junto con la carga mental. Mejores condiciones biomecánicas, también generan una reducción del estrés psicológico.
- En lo que respecta a los cobots, las capacidades más estudiadas han sido la seguridad y la conectividad. Estos dos factores son de particular importancia

en los sistemas colaborativos, ya que están directamente relacionados con la seguridad de la persona operaria y el rendimiento de la celda de trabajo. Por otro lado, la actuación ha sido abordada en menos trabajos, a pesar de su relación con la seguridad de la persona operaria. Resulta interesante observar cómo el paradigma de los robots industriales tradicionales sigue presente en los cobots, ya que la movilidad no se ha considerado un tema principal. Esto es razonable, dado que muchas tareas robóticas requieren una alta precisión que solo se puede lograr mediante una calibración adecuada del espacio de trabajo, un proceso que consume tiempo cada vez que se mueve el robot.

- En cuanto a los sistemas de producción modernos, las características más estudiadas son la orientación al costo y la flexibilidad, directamente derivadas del entorno industrial. Sin embargo, la agilidad, un enfoque fundamental en la Industria 4.0, ha sido escasamente considerada en la investigación en comparación con otras características. Según Faccio et al. (2022), la demanda de productos altamente personalizados en lotes pequeños debería ser de gran interés para futuras investigaciones y podría ser abordada con mayor profundidad.

2.3.2 Herramientas

A continuación, se presentan y analizan las cuatro herramientas, que son:

- i) *Creation of a UX index to design human tasks and workstations* (Grandi et al., 2022)
- ii) *A comprehensive UX index to evaluate industrial tasks from a human-centered perspective* (Khamaisi, Grandi, et al., 2022)
- iii) *UX assessment strategy to identify potential stressful conditions for workers* (Brunzini et al., 2022)
- iv) *On the quantification of human-robot collaboration quality* (Kokotinis et al., 2023)

2.3.2.1 *Creation of a UX index to design human tasks and workstations (Grandi et al., 2022)*

El estudio de Grandi et al. (2022) propone la creación de un Índice de Experiencia de Usuario (UXI) para evaluar la calidad de la interacción humano-sistema durante las tareas laborales y, en consecuencia, evaluar tanto el proceso como el espacio de trabajo. Integra cuatro clases de datos para objetivar la UX y la carga de trabajo percibida: análisis postural, datos fisiológicos, rendimiento y datos subjetivos. El análisis postural se centra en la carga física, los datos fisiológicos se enfocan en la evaluación de la carga mental, los datos de rendimiento brindan una medida de la eficiencia y efectividad de la interacción considerando el tiempo de ejecución, y finalmente los datos subjetivos proporcionan una medida individual de la carga percibida utilizando el cuestionario NASA-TLX.

- **Análisis postural:** Se basa en el método REBA (Hignett & McAtamney, 2000), que proporciona un indicador sintético para evaluar rápidamente el riesgo de trastornos musculoesqueléticos relacionados con el trabajo asociados con la ejecución de ciertas tareas laborales. Este método utiliza un proceso sistemático para evaluar tanto las partes superiores como inferiores del cuerpo para el análisis biomecánico y de riesgos. A partir del análisis de datos posturales, se calcula el parámetro de puntuación REBA para cada usuario, considerando el valor promedio de la puntuación REBA durante la simulación de la tarea (REBA de la persona usuaria), mientras que se considera el valor mínimo (REBA mínimo) como línea de base. Para normalizar este parámetro, se extrae el rango REBA del valor máximo durante la simulación de ensamblaje

para cada usuario, para calcular el parámetro de Confort Postural (PC) de la siguiente manera:

$$PC = \frac{REBA_{usuario} - REBA_{min}}{REBA_{max} - REBA_{min}} \quad (1)$$

- **Análisis fisiológico:** Se lleva a cabo mediante la combinación de diferentes parámetros, como la actividad del ECG, como la frecuencia cardíaca (HR) y la variabilidad de la frecuencia cardíaca (HRV), y el análisis de la actividad ocular, como el diámetro pupilar (PD). Antes de monitorear los datos, se requiere una fase de descanso de aproximadamente 5 minutos para calcular los valores de referencia de los parámetros fisiológicos. Durante esta fase, se les pide a las personas usuarias que vean un video relajante, de pie en una posición cómoda. Los valores de referencia son útiles para calcular parámetros específicos. A partir del análisis de datos fisiológicos, se definen dos parámetros como Actividad Cardíaca (HA) y Actividad Pupilar (PA).

$$HA = \frac{HR_{usuario} - HR_{ref}}{HR_{max} - HR_{ref}} \quad (2)$$

Donde $HR_{usuario}$ es el valor promedio de la frecuencia cardíaca de la persona usuaria específico registrado durante la simulación de la tarea, HR_{ref} es el valor promedio de HR registrado durante la fase de referencia de la persona usuaria, y HR_{max} es el valor máximo de HR registrado para cada usuario durante toda la prueba.

De manera similar, PA se calcula de la siguiente manera:

$$PA = \frac{PD_{usuario} - PD_{ref}}{PD_{max} - PD_{ref}} \quad (3)$$

Donde $PD_{usuario}$ se define como el valor promedio del diámetro pupilar de la persona usuaria específico registrado durante la simulación de la tarea, PD_{ref} es el valor promedio registrado durante la fase de referencia de la persona usuaria, y PD_{max} es el valor máximo registrado para cada usuario durante toda la prueba.

- **Análisis de desempeño:** Se basa en registrar el tiempo necesario para completar la tarea, lo que evalúa el rendimiento de la persona operaria. Se calcula el parámetro Tiempo del Usuario (UT) considerando el rendimiento en términos de tiempo de la persona usuaria. El tiempo necesario para completar la tarea se registra para cada usuario ($T_{usuario}$) y se compara con el tiempo experto ($T_{experto}$) logrado por un operador de ensamblaje experimentado y el tiempo más largo realizado por la persona usuaria con menos experiencia que realizó la prueba (T_{max}), como se muestra en la siguiente ecuación:

$$UT = \frac{T_{usuario} - T_{experto}}{T_{max} - T_{experto}} \quad (4)$$

- **Análisis subjetivo:** se recopilan datos subjetivos a través de un cuestionario posterior basado en el NASA-TLX, que proporciona una evaluación subjetiva de la carga percibida de trabajo. Siendo así, Grandi et al. (2022) consideran las valoraciones de Demanda Física para ponderar el parámetro PC, la valoración de Demanda Mental para ponderar el parámetro PD, la valoración del Nivel de Frustración para ponderar el parámetro HA y, finalmente, la valoración del

Rendimiento General para ponderar el parámetro UT. La suma de estos parámetros, ponderados adecuadamente como se describe, genera la Carga de Trabajo Ponderada (WWL) que suma aspectos tanto físicos como cognitivos, como se muestra en la ecuación (5):

$$WWL = \omega_1 * HA + \omega_2 * PA + \omega_3 * PC + \omega_4 * UT \quad (5)$$

Donde, para cada individuo, ω_1 es el peso relacionado con HA como valoración normalizada para el Nivel de Frustración del cuestionario NASA-TLX, ω_2 es el peso relacionado con PA como valoración normalizada para la Demanda Mental del cuestionario NASA-TLX, ω_3 es el peso relacionado con PC como valoración normalizada para la Demanda Física del cuestionario NASA-TLX, y ω_4 es el peso relacionado con UT como valoración normalizada para el Rendimiento General del cuestionario NASA-TLX.

De este modo, el puntaje total del UXI se calcula en porcentaje de la siguiente manera:

$$UXI = 100\% - \left(\frac{WWL}{20}\right) * 100 \quad (6)$$

Además, Grandi et al. (2022) establecen una escala de valores objetivo para el UXI como punto de referencia para evaluar si los parámetros medidos pueden garantizar una UX positiva. Según esta escala, si el valor del UXI es superior al 80%, la experiencia de usuario se considera positiva. En el rango del 40% al 80%, existe margen para mejorar el diseño y lograr un nivel de comodidad más cercano al óptimo. Sin embargo, si el valor del UXI es inferior al 40%, la comodidad en general se ve comprometida, lo que conlleva un riesgo de sobrecarga tanto física como cognitiva.

2.3.2.2 *A comprehensive UX index to evaluate industrial tasks from a human-centered perspective (Khamaisi, Grandi, et al., 2022)*

El estudio de Khamaisi, Grandi, et al. (2022) tiene como objetivo refinar el índice de UX propuesto por Grandi et al. (2022). Los parámetros considerados para el cálculo del UXI son:

- **Carga Mental de Trabajo (MWL):** Surge de la interacción entre los requisitos de una tarea, las circunstancias en las que se realiza y las habilidades, comportamientos y percepciones de la persona operaria.
- **Estrés (S):** describe la respuesta fisiológica del sistema nervioso simpático.
- **Carga Postural de Trabajo (PWL):** expresada por el puntaje RULA ("Evaluación Rápida de Miembros Superiores"), es un índice ergonómico que evalúa la postura adoptada por la persona usuaria, identificando problemas relacionados con el sistema musculoesquelético.

En la evaluación de la MWL, que proporciona un análisis de la carga mental asociada a la actividad simulada específica, se consideran tres variables principales: la Actividad de la Pupila (PA), la Actividad Cardíaca (HA) y la actividad electrodermal (EDA). Estas variables analizan el cambio en el diámetro de la pupila, el número de latidos cardíacos por unidad de tiempo y la variación de la conductancia de la piel. Todas las variables involucradas se obtienen mediante el postprocesamiento de los datos recopilados por los dispositivos. Cada parámetro fisiológico se basa en la definición de una escala de evaluación única para las diversas actividades. Cabe señalar que la normalización de datos considera una escala de variación entre el valor máximo y mínimo de la variable específica registrada durante toda la ejecución de la tarea: luego, se calcula la variación del valor promedio registrado durante la ejecución de la tarea desde el mínimo. Las fórmulas utilizadas para la normalización se muestran a continuación.

$$HA_{tarea} = \frac{HR_{media} - HR_{min}}{HR_{max} - HR_{min}} \quad (7)$$

$$PA_{tarea} = \frac{PD_{media} - PD_{min}}{PD_{max} - PD_{min}} \quad (8)$$

$$EDA_{tarea, MWL} = \frac{EDA_{media} - EDA_{min}}{EDA_{max} - EDA_{min}} \quad (9)$$

HR_{media} es el valor promedio de la frecuencia cardíaca de la persona usuaria recopilado durante la ejecución de la tarea, HR_{min} y HR_{max} son respectivamente los valores mínimo y máximo de la frecuencia cardíaca. De manera similar, se calculan los parámetros PA y EDA. La normalización según el valor mínimo permite evaluar cuánto se desvía la persona usuaria, en referencia a su condición promedio, de la condición de esfuerzo mínimo dentro de la tarea específica.

Así, el parámetro MWL se calcula según:

$$MWL = HA_{tarea} + PA_{tarea} + EDA_{tarea, MWL} \leq 3 \quad (10)$$

La evaluación del estrés se basa en considerar dos parámetros fisiológicos: EDA e IBI ("Intervalo entre Latidos", índice que define el intervalo de tiempo entre dos latidos cardíacos sucesivos). Ambos valores se calcularon considerando la diferencia, en valor absoluto, entre el valor promedio registrado durante la tarea y el valor promedio de referencia (es decir, el valor registrado durante una sesión de referencia de 3 minutos en una posición de reposo). Se asumió que la sesión de referencia representaba la condición libre de estrés para la persona operaria.

El parámetro S se calcula:

$$S_{tarea} = EDA_{tarea, s} + IBI_{tarea} \leq 2 \quad (11)$$

Donde:

$$EDA_{tarea, s} = |EDA_{media, tarea} - EDA_{media, ref}| \quad (12)$$

$$IBI_{tarea, s} = |IBI_{media, tarea} - IBI_{media, ref}| \quad (13)$$

En lo que respecta a la Carga Postural de Trabajo (PWL), se considera el valor promedio del puntaje RULA durante toda la tarea. Específicamente, dado que el puntaje RULA se puede calcular por separado para la parte derecha e izquierda del cuerpo humano, se considera el valor máximo entre los dos para cada momento en el tiempo: de esta manera, se determina un único puntaje representativo de la peor condición postural. En resumen, la PWL se define como en (13):

$$PWL = media (RULA) \leq 7 \quad (14)$$

Dado que la MWL devuelve un puntaje entre 0 y 3, el Estrés devuelve un valor entre 0 y 2, y el RULA, por definición, está en el rango entre 1 y 7, se podría deducir que el UXI asume valores entre 1 y 12. Esto implica que cuanto mayor sea el valor del UXI, más la tarea asociada presenta problemas críticos que deben ser analizados, siguiendo los pasos que llevaron a la definición del índice en sí mismo, e identificando qué parámetros influyen en el rendimiento general.

2.3.2.3 *UX assessment strategy to identify potential stressful conditions for workers (Brunzini et al., 2022)*

El estudio de Khamaisi, Brunzini, et al. (2022) presentan una estrategia para evaluar la UX de las personas trabajadoras y promover una visión centrada en el ser humano de los sitios de fabricación, mejorando la sostenibilidad general de las fábricas modernas. Se utiliza un conjunto de dispositivos portátiles no invasivos para monitorear las actividades humanas y recopilar parámetros fisiológicos, así como cuestionarios para recopilar autoevaluaciones subjetivas.

El enfoque metodológico presentado por Khamaisi, Brunzini, et al. (2022) tiene como objetivo crear un estándar para la evaluación de la UX para respaldar el diseño de todo tipo de procesos industriales o productos y, en general, del entorno de trabajo. Se basa en una estrategia de mapeo de UX, como una secuencia de actividades capaces de analizar la UX de las personas operarias y proponer sugerencias para su optimización. Por lo tanto, debe especificarse que este enfoque atraviesa la aplicación específica elegida y podría adoptarse en varias áreas de aplicación y en diferentes entornos (tanto simulados como en el campo).

Tal y como se observa en la Tabla 25, la primera fase implica una investigación exhaustiva sobre el contexto objetivo y, específicamente, sus características principales. Tal análisis puede realizarse mediante entrevistas o encuestas guiadas a expertos en procesos de la empresa. La Tabla 25 muestra los aspectos a analizar. Este análisis es necesario para identificar las características que podrían hacer que cada tarea sea estresante, riesgosa físicamente, sobrecargada mentalmente o peligrosa para el bienestar de la persona operaria.

Posteriormente, se lleva a cabo una investigación de UX ("Análisis de la UX percibida por las personas operarias") para ampliar la carga de trabajo percibida por la persona operaria desde un punto de vista cognitivo y físico. En particular, se proporcionan cuestionarios de autoevaluación subjetiva a las personas operarias involucrados en diferentes tareas. Después de la ejecución de la tarea, se les pide que completen encuestas específicas para recopilar datos demográficos y su retroalimentación sobre sus roles en el trabajo, la carga de trabajo percibida, la experiencia de la persona usuaria, la comodidad, el malestar en diferentes partes del cuerpo o su familiaridad con la tecnología y dispositivos innovadores que podrían adoptarse para apoyar el proceso (como herramientas para experiencias de realidad extendida o dispositivos portátiles).

Esta evaluación de doble enfoque basada en tareas y operadores es fundamental para la "Definición del objetivo de optimización de la UX". De hecho, el objetivo final de la evaluación de la UX es identificar y resolver los problemas planteados por tarea. La optimización de la UX tiene como objetivo mejorar la ergonomía física, cognitiva, ambiental, organizativa o una combinación de algunas de ellas. Esto influye inevitablemente en la elección de la población de prueba final para monitorear y en la que basar inferencias: en general, para reforzar las consideraciones resultantes, los sujetos involucrados deberían al menos reflejar a la población analizada, tanto cuantitativa como cualitativamente.

Después de la definición del objetivo, es necesario determinar las variables a monitorear ("Definición de las variables a monitorear"). Según los objetivos de la evaluación de la UX, pueden surgir varias configuraciones de variables a monitorear. Luego, según las

variables seleccionadas y la configuración del lugar de trabajo y las restricciones, se pueden seleccionar el *hardware* y el *software* para la evaluación de la UX ("Definición de la configuración tecnológica"). Basándose en las condiciones límite (por ejemplo, esfuerzo físico predominante en lugar del cognitivo, o la posibilidad de investigar las percepciones de la persona operaria en tiempo real), es necesario definir el algoritmo para la evaluación de la UX ("Definición del algoritmo").

Cuando se definen todos estos pasos, puede llevarse a cabo la "Sesión de Monitoreo" y completarse la "Evaluación de la UX". La sesión de monitoreo consiste en la adopción de la configuración seleccionada para recopilar datos humanos durante la ejecución de la tarea, en diferentes personas usuarias, con el fin de aplicar la estrategia de mapeo de UX propuesta. Posteriormente, la evaluación de la UX se basa en el procesamiento posterior de los datos y su interpretación.

Tabla 25: Flujo de trabajo de la estrategia de mapeo UX. Fuente: Khamaisi, Brunzini, et al. (2022).

Fase	Descripción	Aspectos a analizar
Fase 1	Análisis de la organización del trabajo	Turnos de trabajo (incluyendo turnos nocturnos), planificación de pausas, relaciones laborales, formación, carga cognitiva, tipos de interacciones laborales, nivel de esfuerzo físico, entorno de trabajo, equipo y herramientas utilizadas.
Fase 2	Análisis de la UX percibida por las personas operarias	Recopilación de comentarios de las personas operarias, cuestionarios de experiencia de la persona usuaria, evaluación de la incomodidad en diferentes partes del cuerpo, cuestionario de autoeficacia tecnológica, entre otros.
Fase 3	Definición del objetivo de la optimización de la UX	Identificación de aspectos de ergonomía física, cognitiva, medioambiental y organizativa, además de la consideración de la experiencia de la persona usuaria en general.
Fase 4	Definición de las variables a monitorizar	Parámetros relevantes a monitorear, como manipulación manual, repetición de tareas, calidad del aire, polvo, contaminación, temperatura, malestar corporal percibido, estrés percibido y esfuerzo mental, así como parámetros fisiológicos.
Fase 5	Definición de la implantación tecnológica	Elección de la tecnología para la implementación, ya sea a través de simulaciones de realidad virtual (<i>hardware</i> y <i>software</i>) o aplicaciones en el entorno de trabajo real (<i>hardware</i> y <i>software</i>).
Fase 6	Definición del algoritmo	Establecimiento de condiciones límite basadas en el esfuerzo físico predominante, el esfuerzo cognitivo predominante y la capacidad de investigar las percepciones en tiempo real de las personas operarias.
Fase 7	Monitorización de la sesión	Registro de datos en tiempo real durante las sesiones de trabajo, incluyendo parámetros fisiológicos, interacciones, respuestas de las personas operarias, etc.
Fase 8	Evaluación de la UX	Evaluación de los datos recopilados para determinar la experiencia de la persona usuaria, identificación de puntos problemáticos y áreas de mejora, así como comparación con objetivos de optimización establecidos.

2.3.2.4 On the quantification of human-robot collaboration quality (Kokotinis et al., 2023)

El estudio de Kokotinis et al. (2023) propone un enfoque para evaluar la calidad de la colaboración en entornos de trabajo de HRC. El objetivo principal es proporcionar un índice medible que se calcule en base a indicadores individuales que reflejen los aspectos de rendimiento, costo, calidad y flexibilidad de la aplicación de HRC. Este índice implica la cuantificación de métricas pertenecientes a tres categorías diferentes, como se muestra en la Figura 22.

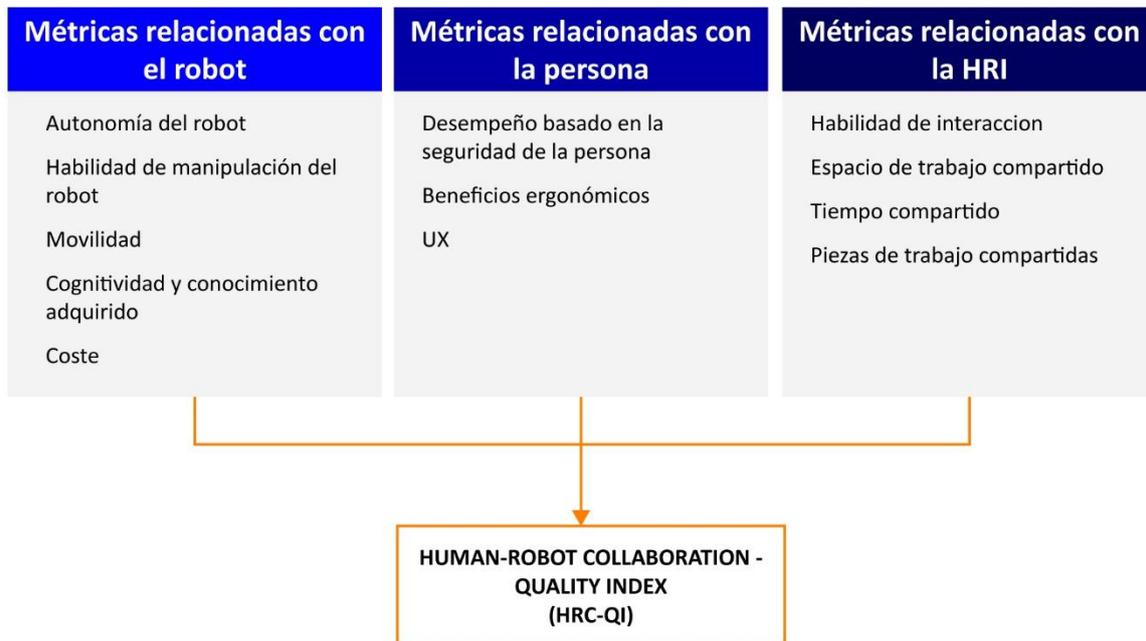


Figura 22: Índice de Calidad para medir la HRC. Adaptado de Kokotinis et al. (2023).

Según los autores, estas tres categorías de métricas se componen de:

Métricas relacionadas con el robot: Estas métricas buscan capturar la contribución de los recursos robóticos a la colaboración, lo cual tiene un impacto directo en la aceptación por parte de las personas operarias y en los beneficios de productividad que se pueden lograr. Los robots que exhiben autonomía, inteligencia, destreza en la manipulación, entre otros aspectos, pueden colaborar de manera más eficiente con las personas y generar una aplicación de HRC de mayor rendimiento. Las métricas incluidas en esta categoría son:

- **Autonomía del robot:** define la capacidad de un recurso robótico para trabajar de forma autónoma en un entorno colaborativo sin necesidad de la intervención de la persona operaria. A mayor autonomía, menor es el esfuerzo requerido por parte de la persona en el contexto de la colaboración.
- **Capacidad de manipulación del robot:** refleja la habilidad de un sistema robótico para agarrar y manejar un objeto dentro del contexto de una tarea. Cuanto más capaz sea el robot, más tareas puede realizar y contribuir de diversas formas a la colaboración.
- **Movilidad:** se refiere a la capacidad del recurso robótico para desplazarse o ser reposicionado manualmente en toda el área de producción con el fin de ejecutar tareas en diferentes estaciones de trabajo. Esta habilidad mejora la accesibilidad del robot y, por lo tanto, su potencial para trabajar más cerca de las personas. En combinación con la capacidad de aprender y ejecutar autónomamente nuevas tareas, este atributo mejora considerablemente las aplicaciones colaborativas que el robot puede abordar.
- **Capacidad de cognición y adquisición de conocimiento:** es la capacidad de un robot para adquirir conocimiento sobre su entorno y objetos, y utilizarlo para ajustar

dinámicamente su comportamiento a lo largo del tiempo. Cuando se combina con la evaluación de la ergonomía humana, puede mejorar el apoyo que las personas reciben a medida que la colaboración se extiende en el tiempo.

- Costo: se refiere al costo de un recurso robótico en comparación con el costo correspondiente necesario para que una persona operaria realice la misma operación.
- Aunque esto no afecta directamente la calidad de la colaboración, es uno de los principales factores a considerar al invertir en cualquier aplicación.

Métricas relacionadas con la persona: Esta categoría engloba métricas que cuantifican la experiencia percibida por las personas operarias en términos de seguridad humana, ergonomía, escepticismo o confianza en la colaboración con robots, así como la experiencia general de la persona usuaria. Básicamente, cualquier métrica en esta categoría debe cuantificar el impacto directo de la aplicación de HRC en la persona operaria y su perspectiva hacia dicha aplicación, ya sea de manera positiva o negativa. Las métricas incluidas en esta categoría son:

- Rendimiento para la seguridad humana: esta métrica considera la cantidad y el tipo de funciones de seguridad presentes en la aplicación HRC, las cuales pueden permitir o prohibir que el robot trabaje a alta velocidad. Expresa el equilibrio entre la productividad y las limitaciones de velocidad/rendimiento necesarias para garantizar la seguridad de las personas.
- Beneficios ergonómicos: representa el apoyo físico y/o mental que una persona trabajadora puede recibir al trabajar en una aplicación colaborativa específica. Cuanto mayor sea el nivel de esta métrica, más apoyo recibirán las personas operarias del sistema integrado y, por lo tanto, más beneficios obtendrán en términos de estrés y fatiga a largo plazo.
- UX: refleja las preferencias personales de las personas operarias y la sensación general que desarrollan al trabajar con robots. También considera el nivel de confianza o miedo/escepticismo que las personas pueden desarrollar mientras trabajan o interactúan con cualquier sistema robótico.

Métricas relacionadas con la HRI: Esta categoría de métricas se centra en cuantificar la experiencia de interacción disponible a través de las diferentes modalidades que los robots pueden utilizar para interactuar y coexistir con las personas. Los sistemas sensoriales de los robots, los dispositivos periféricos y las interfaces hombre-máquina son elementos que habilitan diferentes esquemas de colaboración, como el intercambio de espacio de trabajo, tiempo y piezas de trabajo. Estos esquemas tienen un efecto directo en la calidad percibida de la colaboración, ya que influyen en la forma en que se pueden compartir las tareas entre personas y robots. Las métricas incluidas en esta categoría son:

- Capacidad de interacción: considera las diferentes formas de intercambiar información con un robot. El número de opciones disponibles, así como el uso intuitivo de cada opción, crea una experiencia diferente para la persona usuaria, quien puede encontrar más o menos fácil trabajar con el sistema.
- Compartir espacio de trabajo: denota la capacidad de la persona operaria y el recurso robótico para ejecutar operaciones en el mismo espacio de trabajo, independientemente del momento o las piezas que se estén procesando. Un sistema robótico tradicional con vallas permite compartir el espacio de trabajo en comparación con una celda robótica completamente colaborativa sin separación física.
- Compartir tiempo: representa la capacidad de la persona operaria y el recurso robótico para trabajar en la misma tarea de forma paralela.
- Piezas de trabajo compartidas:

Con el fin de obtener una evaluación sencilla de cada métrica, Kokotinis et al. (2023) seleccionaron un conjunto de niveles considerando los casos más representativos (Tabla 26). El nivel/valor más bajo implica una HRC no eficiente, mientras que el nivel más alto corresponde al escenario ideal en el que el robot se comporta como un segundo ser humano. Por supuesto, en algunos casos no es posible alcanzar el escenario "ideal", sin embargo, es un buen punto de referencia para la comparación entre aplicaciones similares.

Tabla 26: Evaluación para medir el HRC-QI. Adaptado de: Kokotinis et al. (2023).

Categoría	Métrica	Descripción del escenario	Valor de la métrica
Métricas relacionadas con el robot	Autonomía del robot	Entendimiento del flujo de trabajo y ejecución de acciones sin intervención humana.	5
		Autonomía supervisada en términos de ejecutar una tarea de forma autónoma en la mayoría de las condiciones de operación y reconocer cuándo se necesita intervención humana.	4
		Autonomía compartida en términos de ejecución parcial del flujo de trabajo de forma autónoma y la intervención humana para seleccionar tareas posteriores.	3
		Ejecución de acciones predefinidas de forma semiautónoma mientras la persona interviene para controlar el sistema y llevar a cabo una tarea.	2
		Requiere instrucción/programación/confirmación en cada paso.	1
	Capacidad de manipulación del robot:	Agarrar/manipular cualquier parte requerida para el proceso.	5
		Utilizar cambiadores de herramientas, permitiendo el uso de múltiples efectores finales que pueden manejar diferentes tipos/tamaños/pesos de productos.	4
		Agarrar/manipular solo partes dentro de un rango específico de tamaño/peso/flexibilidad utilizando un único efector final, pero "flexible".	3
		Agarrar/manipular un número limitado de componentes utilizando un efector final dedicado a una parte específica (por ejemplo, una ventosa).	2
		Agarrar/manipular solo partes con geometría y tamaño específicos (por ejemplo, destornillador para un tamaño específico de tornillos).	1
	Movilidad	Un robot es capaz de desplazarse o ser reubicado dentro de todo el lugar de trabajo para ejecutar una tarea (por ejemplo, un manipulador móvil).	3
		Un robot es ligero y puede ser movido por la persona.	2
		El robot es estacionario.	1
	Capacidad de cognición y adquisición de conocimiento	Aprendizaje por observación adquiriendo conocimiento indirectamente a través de la observación de otros robots o personas llevando a cabo tareas.	6
		Conocimiento de la persona usuaria adquiriendo conocimiento sobre la operación al observar a las personas en un área definida.	5
		Conocimiento solicitado solicitando información humana sobre un objeto o lugar relevante para la tarea.	4
		Conocimiento del lugar relacionado con la ubicación/tipo de objetos y características ambientales mediante la coincidencia de objetos con tipos predefinidos y conocidos.	3
		Conocimiento de datos del sensor - acumulando conocimiento sobre el entorno basado en datos de sensores.	2
		No se adquiere conocimiento por parte del robot durante su operación.	1

Métricas relacionadas con la persona	Coste	Menor que un empleado que realiza las mismas tareas.	3
		Igual o un poco más alto que un empleado que realiza las mismas tareas.	2
		Significativamente más alto que un empleado que realiza las mismas tareas.	1
	Rendimiento para la seguridad humana	Seguridad intrínseca en términos de que el robot es seguro por diseño y no puede ejercer una fuerza que sea dañina para la persona o trabajar a una velocidad peligrosa para las personas.	5
		Trabajar al 100% de la velocidad diseñada en todo momento, incluso cuando hay una persona presente, percibir el entorno y realizar todas las funciones de seguridad requeridas (por ejemplo, evitar colisiones, limitación de potencia y fuerza, etc.), pero reconocer cuando esta forma de trabajo pueda poner en peligro al usuario y modificar su movimiento/velocidad para garantizar una operación segura.	4
		Seguridad adaptada mediante el robot trabajando a una velocidad reducida diseñada para tareas colaborativas específicas en todo momento, sin necesariamente significar que la persona esté a una distancia cercana.	3
		Seguridad básica en términos de que el robot reconoce que una persona se acerca/inserta físicamente en su espacio de operación y detiene su operación automáticamente.	2
		Sin seguridad: el robot no está equipado para trabajar con una persona.	1
	Beneficios ergonómicos	Soporte físico completo (partes pesadas/peligrosas) y soporte mental, reduciendo el estrés y la fatiga humana en el espacio de trabajo de operaciones peligrosas y repetitivas o proporcionando indicaciones sobre el estado del flujo de trabajo.	5
		Soporte físico limitado (por ejemplo, co-manipulación: compartir peso), pero en combinación con soporte mental durante condiciones peligrosas.	4
		Solo soporte físico en aplicaciones colaborativas, donde el robot es responsable de manejar partes pesadas/peligrosas.	3
		Solo soporte mental, donde el robot proporciona información o indicaciones sobre el estado del flujo de trabajo (visualmente, auditivamente u otros) o reduce el estrés de las operaciones repetitivas.	2
		Sin beneficio ergonómico.	1
	Experiencia de Usuario	Sensación de seguridad en la HRC y plena confianza en todo momento.	3
		HRC segura certificada, pero confianza limitada (por ejemplo, persona intimidada por un robot colaborativo de alta carga útil).	2
Miedo y escepticismo de coexistir con un robot (por ejemplo, un robot sobre la cabeza de la persona) - falta de voluntad para cooperar.		1	
Métricas relacionadas con la interacción	Capacidad de interacción	Todos los tipos de interacción y funcionalidades están disponibles (directas, indirectas) e implementadas de manera muy eficiente, suave y natural, permitiendo también el control indirecto del robot (interacción basada en gestos para mover el robot).	6
		Capacidad de interacción directa completa combinada con interacción indirecta solo para funciones básicas de iniciar/detener el robot y visualizar información relacionada con el robot (por ejemplo, a través de herramientas basadas en realidad aumentada).	5
		Interacción física directa en términos de enseñar una secuencia de movimiento al robot a través de retroalimentación háptica simple (contacto).	4
		Control directo evaluando el estado del robot a través de	3

	retroalimentación visual/visual y proporcionando control del robot momento a momento (por ejemplo, usando un joystick).	
	Capacidad de interacción indirecta limitada en términos de iniciar/detener el robot a través de un dispositivo de control (por ejemplo, un reloj inteligente o una interfaz hombre-máquina).	2
	No hay interacción disponible.	1
Espacio de trabajo compartido	Capacidad de trabajar en todas las áreas como su contraparte humana sin restricciones.	5
	Espacio de trabajo parcialmente compartido.	4
	Mismo espacio de trabajo, pero trabajando en tareas diferentes.	3
	Mismo espacio de trabajo, pero se permite que la persona ingrese cuando el robot se detiene y viceversa.	2
	Espacio de trabajo completamente separado.	1
Tiempo compartido	Capacidad de trabajar en paralelo.	3
	Trabajo consecutivo: cada entidad espera a que la otra termine.	2
	La persona realiza todas sus tareas y luego el robot comienza a realizar las suyas.	1
Piezas de trabajo compartidas	Compartir completamente la pieza de trabajo en términos de que humanos y robots usen las mismas partes/herramientas, si es necesario, para ejecutar el proceso.	4
	Compartir pieza de trabajo con interferencia mínima de los volúmenes de trabajo del robot/humano.	3
	Compartir pieza de trabajo sin interferencia de los volúmenes de trabajo del robot/humano, manteniendo una distancia segura en todo momento (por ejemplo, diferentes lados de una pieza larga).	2
	Sin compartir: solo una persona o un robot trabaja en la pieza al mismo tiempo.	1

2.3.3 Estudios experimentales

A continuación, se presentan y analizan los estudios experimentales identificados, que son:

- i) *Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures* (Akalin et al., 2022)
- ii) *User Experience and Physiological Response in Human-Robot Collaboration: A Preliminary Investigation* (Gervasi et al., 2022)
- iii) *Manual assembly and Human–Robot Collaboration in repetitive assembly processes: a structured comparison based on human-centered performances* (Gervasi et al., 2023)

2.3.3.1 *Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures* (Akalin et al., 2022)

Akalin et al. (2022) investigaron la relación entre la comodidad, la previsibilidad, el sentido de control y la confianza con la seguridad percibida en la HRI utilizando medidas objetivas y subjetivas. Llevaron a cabo un experimento de diseño mixto con un esquema de dos por cinco. Los investigadores establecieron dos condiciones entre sujetos: la interacción con el robot defectuoso se realizó al principio o al final del experimento. Además, definieron cinco condiciones dentro de los sujetos que abarcaron desde la línea de base hasta manipulaciones específicas del comportamiento del robot para inducir situaciones de incomodidad, disminución de la seguridad percibida, disminución del sentido de control y desconfianza.

La metodología del estudio incluyó la recopilación de datos tanto subjetivos como objetivos. Las personas participantes (n=27) respondieron cuestionarios después de cada una de las condiciones dentro de los sujetos, evaluando factores específicos relacionados con la experiencia. También se recopilaron medidas objetivas como videos y datos fisiológicos.

Las cinco condiciones intra-sujeto del experimento fueron:

- Situación inicial (C1)
- Manipulación cómoda (C2)
- Comportamientos impredecibles del robot (C3)
- Manipulación de la sensación de control (C4)
- Manipulación de la confianza (C5)

Estas condiciones se organizaron en dos procedimientos; A y B. La única diferencia entre estas dos era el orden de las condiciones. En la Configuración A, C3 y C5 ocurrieron después de las otras condiciones, el orden fue el siguiente: C1, C4, C2, C3 y C5. En la Configuración B, C3 y C5 ocurrieron después de la condición de referencia, el orden fue el siguiente: C1, C3, C5, C4 y C2.

Las personas participantes completaron un cuestionario previo al experimento, cuestionarios entre condiciones y un cuestionario posterior al experimento. Para empezar, utilizaron un cuestionario de personalidad basado en el *Big Five Inventory* (Rammstedt & John, 2007). Los cuestionarios entre experimentos fueron un cuestionario sobre el confort basado en el estudio de (Kim & Mutlu, 2014), un cuestionario sobre la seguridad percibida (Akalin et al., 2019), un cuestionario sobre el sentido de control adaptado de Strube & Werner, (1984) y el cuestionario *Trust Perception Scale-HRI* (Schaefer, 2016). El cuestionario post ejecución fue el de Weiss et al. (2009).

Además, utilizaron un análisis de expresiones faciales, concretamente Affdex SDK (McDuff et al., 2016). Affdex produce un conjunto de características que incluyen siete emociones (enojo, desprecio, disgusto, miedo, alegría, tristeza y sorpresa), compromiso (expresividad facial del participante), valencia (la agradable sensación del participante), 20 expresiones faciales (fruncir el ceño, levantar la ceja, levantar el cachete, levantar la barbilla, hoyuelos, cerrar los ojos, abrir los ojos, levantar la ceja interna, dejar caer la mandíbula, apretar los párpados, deprimir las esquinas de los labios, presionar los labios, fruncir los labios, estirar los labios, chupar los labios, abrir la boca, arrugar la nariz, sonreír, sonreír con malicia y levantar el labio superior), y atención (basada en la orientación de la cabeza).

Adicionalmente, recopilaron fisiológicos utilizando la pulsera Empatica E4 , que mide el Pulso de Volumen Sanguíneo (PVS), el Acelerómetro de 3 ejes (ACC), la EDA, la temperatura periférica de la piel (ST) y la Frecuencia Cardíaca (HR), con las siguientes frecuencias de muestreo: 64 Hz, 32 Hz, 4 Hz, 4 Hz y 1 Hz, respectivamente. La pulsera también proporciona los intervalos de latidos intercardíacos (IBI), los cuales no tienen una frecuencia de muestreo específica.

El robot utilizado durante la experimentación fue el robot humanoide Pepper (Pandey & Gelin, 2018). Se trata de un robot social que admite la comunicación bidireccional con lenguaje natural a través de un *software* de texto a voz.

Tras analizar sus resultados, los principales resultados y directrices para aumentar la percepción de seguridad en la HRI según Akalin et al. (2022) son:

- Se debe enfocar en comprender las condiciones en las que las personas se sienten inseguros en lugar de cuando se sienten seguros. Las medidas cuantificables surgen en condiciones inseguras.

- En relación a las medidas objetivas y subjetivas, los factores relacionados con el robot y los factores relacionados con las personas deben ser abordados conjuntamente debido a la naturaleza bidireccional de la IPR.
- Los principales factores influyentes en la percepción de seguridad son la comodidad, la experiencia/familiaridad, la predictibilidad, el sentido de control, la transparencia y la confianza.
- Estos factores deben ser considerados en las decisiones de diseño de la IPR para garantizar la seguridad. Los factores relacionados con el robot no deben resultar en incomodidad, falta de control y desconfianza por parte de las personas usuarias. Además, los comportamientos del robot deben ser familiares, predecibles y transparentes.
- Los resultados indican que la tasa de predicción de la percepción de seguridad es mayor cuando se utilizan datos de señales fisiológicas.
- Finalmente, se debe tener en cuenta las características individuales de las personas, las reacciones emocionales y fisiológicas, así como la interrelación entre los factores para comprender mejor la fuente de la disminución de la percepción de seguridad.

2.3.3.2 User Experience and Physiological Response in Human-Robot Collaboration: A Preliminary Investigation (Gervasi et al., 2022)

El estudio de Gervasi et al. (2022) se enfoca en los aspectos humanos relacionados con la HRC en el entorno industrial, investigando cómo diferentes configuraciones de la interacción influyen en la UX, el estado afectivo y el nivel de estrés. Para esto, se implementó una tarea de ensamblaje colaborativo con variaciones en la velocidad de movimiento del robot, la distancia entre el operador y el espacio de trabajo del robot, y el control del tiempo de ejecución de la tarea.

En el estudio participaron 42 personas, quienes proporcionaron retroalimentación sobre la calidad de la interacción y su estado afectivo. Además, se recopilaron datos fisiológicos de manera no invasiva, como la actividad electrodermal y la frecuencia cardíaca, para monitorear el nivel de estrés durante la interacción.

Asimismo, es relevante destacar el equipamiento utilizado durante la experimentación para recopilar datos precisos. Se empleó la pulsera Empatica E4 biosensor de Empática (2022) para recolectar datos de la EDA a una frecuencia de 4Hz, datos de frecuencia cardíaca a través de la fotopleletismografía (PPG) a 64Hz y datos del acelerómetro de 3 ejes a 32Hz. Esta pulsera también proporcionó intervalos de NN de la frecuencia cardíaca, que fueron útiles para el monitoreo fisiológico de las personas participantes.

Además de la recolección de datos fisiológicos, se utilizaron diversos cuestionarios para capturar diferentes aspectos de la experiencia de las personas participantes. Se aplicó un cuestionario sociodemográfico para recabar información sobre la experiencia previa con cobots y datos demográficos relevantes. El cuestionario NARS (Nomura et al., 2004) se utilizó para evaluar la actitud de las personas participantes hacia los robots. Para medir las reacciones afectivas de las personas participantes basadas en el modelo Placer-Arousal-Dominio (PAD), se empleó la autoevaluación Manikin (SAM). Por último, se utilizó un cuestionario específico sobre la calidad de la interacción para evaluar la percepción de las personas participantes en cuanto a diversos aspectos de la interacción con el cobot.

De este estudio, los autores obtuvieron las siguientes conclusiones:

- La velocidad de movimiento del robot y el control del tiempo de ejecución de la tarea son los factores de configuración más influyentes. Una velocidad baja del robot se relaciona con una percepción reducida de utilidad, eficiencia, naturalidad y fluidez de

la interacción, pero aumenta la sensación de seguridad y comodidad. No se encontraron diferencias significativas en la calidad de la interacción entre velocidades media y alta.

- El control del tiempo de ejecución de la tarea tiene un impacto negativo significativo en la calidad de la interacción y el estado afectivo de las personas participantes. La falta de control sobre el tiempo de ejecución hace que el robot sea percibido como menos útil y confiable, y la interacción menos natural, eficiente, segura, fluida y cómoda. El control del tiempo de ejecución se asocia con una mayor valencia y dominio, así como una menor activación, lo que indica una experiencia más relajada. Además, la falta de control del tiempo de ejecución puede generar estrés fisiológico.
- La distancia entre el operador y el espacio de trabajo del robot no tiene un impacto significativo en la experiencia de las personas participantes, aunque una mayor distancia se relaciona con una ligera mejora en la confianza en el robot. La distancia es principalmente una cuestión de preferencia personal y no influye de manera relevante en la tarea colaborativa.
- Se identifican posibles efectos debido a las interacciones entre la velocidad y la distancia, así como entre la velocidad y el control. La proximidad de la persona operaria al espacio de trabajo del robot tiene efectos diferentes según la velocidad del robot, lo que sugiere que la configuración óptima puede variar en función de estos dos factores. Además, la falta de control del tiempo de ejecución de la tarea tiene un impacto negativo en la experiencia de la persona usuaria, especialmente a altas velocidades.
- A medida que avanzan los ensayos experimentales, se observa una mejora gradual en la experiencia de la persona usuaria. Las dimensiones de la calidad de la interacción y la sensación de control y relajación de las personas participantes mejoran con el tiempo, lo que indica un proceso de habituación.

2.3.3.3 Manual assembly and Human–Robot Collaboration in repetitive assembly processes: a structured comparison based on human-centered performances (Gervasi et al., 2023)

Gervasi et al. (2023) presentan un estudio experimental que investiga los efectos de la HRC prolongada en la UX y el rendimiento en un proceso de ensamblaje repetitivo. El objetivo del estudio fue analizar las diferencias entre un ensamblaje manual y uno realizado en colaboración con un robot en términos de variables relacionadas con el estado afectivo, el malestar corporal, la carga de trabajo y las señales fisiológicas de estrés, proceso y defectos del producto.

Para recopilar los datos, se utilizaron diferentes instrumentos y mediciones. El estado afectivo autoinformado se evaluó utilizando el cuestionario NASA-TLX y el Self-Assessment Manikin (SAM), que permitió recopilar información sobre las dimensiones de demanda mental, demanda física y demanda temporal. Además, se recolectaron datos sobre el malestar corporal percibido, la carga de trabajo percibida y las señales fisiológicas de estrés, como la variabilidad de la frecuencia cardíaca y la actividad electrodermal.

Los resultados del estudio mostraron que la HRC en el proceso de ensamblaje repetitivo tuvo efectos positivos. Específicamente, se observó una mejora en la ergonomía física, una reducción en el esfuerzo mental y el estrés, así como una disminución en los defectos del proceso. Estos hallazgos destacan la capacidad de la robótica colaborativa para mejorar la calidad del proceso al brindar apoyo cognitivo a las personas operarias en situaciones repetitivas.

En términos de implicaciones prácticas, el estudio resalta la importancia de considerar los factores humanos en los procesos de fabricación que involucran la HRC. Esto permitirá

mejorar la calidad y adaptabilidad de los procesos de producción, beneficiando tanto a las personas operarias como a la eficiencia general del sistema.

Gervasi et al. (2023) mencionan también que el estudio presenta ciertas limitaciones. En primer lugar, se realizó con una muestra pequeña, lo que podría limitar la generalización de los resultados a otras poblaciones. Además, el estudio se llevó a cabo en un entorno de laboratorio, lo que plantea interrogantes sobre la aplicabilidad de los resultados en entornos de fabricación reales. Por último, es importante tener en cuenta que el estudio se centró exclusivamente en los efectos de la HRC en un proceso de ensamblaje repetitivo, por lo que sus resultados pueden no ser directamente aplicables a otros tipos de procesos de fabricación.

2.4 Métodos y métricas para evaluar la UX y la aceptación tecnológica

Los métodos de evaluación de la UX y la usabilidad se han convertido en motivo de preocupación. Existe una amplia gama de métodos de evaluación tanto para la UX como para la usabilidad de los sistemas de información. Las dos principales categorías de métodos de evaluación de UX se denominan subjetivas, generalmente medidas a través de entrevistas, encuestas, calificaciones, cuestionarios, etc., y objetivas, generalmente medidas a través de medidas de rendimiento de la persona usuaria (por ejemplo, frecuencia de errores, tiempo de finalización de tareas, etc.), respuestas de atributos fisiológicos y neurológicos (por ejemplo, señales cardiovasculares, ondas cerebrales, seguimiento ocular, expresiones faciales, respuesta de la piel, etc.), evaluación cognitiva, etc. (Mandryk et al., 2006; Robinson et al., 2018). De manera similar, una de las principales categorías de métodos de evaluación de usabilidad es la evaluación sumativa (Zaki et al., 2021), que incluye medidas tanto subjetivas como objetivas, donde las medidas objetivas incluyen pruebas de usabilidad y de referencia utilizando atributos fisiológicos y neurológicos y métricas de usabilidad; y las medidas subjetivas incluyen cuestionarios, entrevistas y encuestas. Estos métodos de evaluación se replican con la participación de personas usuarias, lo que difiere del método de heurísticas que se enfoca en la participación de evaluadores expertos.

Aunque hay métodos que utilizan un enfoque automatizado para la recopilación de datos de la UX (Maia & Furtado, 2016b), hay menos ejemplos de un método bien definido que combine la recopilación automática de datos y la evaluación automática de la UX con las personas. Es por ello, que se ha visto la necesidad de realizar una revisión sistemática de la literatura sobre la evaluación de la UX con herramientas tecnológicas fisiológicas. Además, el continuo desarrollo de sistemas y dispositivos, y la correspondiente necesidad de sistemas más interactivos motivan el dominio de la UX a navegar por las posibles opciones de mejora (Taffese, 2017). Maia & Furtado (2016b) han realizado una revisión sistemática de la literatura sobre la evaluación de la UX. Consideran que el mejor momento para la recopilación de datos es al final de la experiencia o durante y después de ella. También obtienen como resultado que la mayoría de los estudios realizan la recopilación de datos manualmente, algunos lo hacen de forma mixta (manual y automatizada) y muy pocos lo hacen de forma automática. Los resultados obtenidos evidencian que la medición de las emociones de la persona usuaria sigue siendo un desafío para los investigadores del HCI (Maia & Furtado, 2016a).

Robinson et al. (2018) subrayaron la diversidad de métodos utilizados por los investigadores de UX. Sus conclusiones concordaron con investigaciones previas (Bargas-Avila & Hornbæk, 2011) y destacaron la predominancia de técnicas como encuestas, entrevistas, tests de usabilidad y grupos de discusión. Esta inclinación hacia métodos adoptados de otras disciplinas sugiere que la investigación en UX se nutre de diversas aproximaciones. Sin embargo, es digno de mención que la incorporación de medidas fisiológicas ha sido poco frecuente en los estudios de UX (Apraiz et al., 2021).

2.4.1 Evaluación psicofisiológica

En la búsqueda constante por comprender la compleja interacción entre personas y tecnología, la utilización de dispositivos fisiológicos ha emergido como una herramienta fundamental en la evaluación de la UX y la aceptación tecnológica. De acuerdo con Neumann & Westbury (2011), las medidas fisiológicas son más objetivas que las medidas de autorregistro, como los cuestionarios. Las oportunidades ofrecidas por las señales fisiológicas, además, están en aumento debido a la evolución de los sensores y las señales de procesamiento (Oliveira & Guimarães, 2013). Los dispositivos fisiológicos operan en el cruce entre la psicología y la fisiología, y su función principal radica en medir y cuantificar las respuestas que el cuerpo humano experimenta ante diversos estímulos. Estos dispositivos pueden abarcar una variedad de parámetros, desde la frecuencia cardíaca y la conductancia de la piel hasta el electroencefalograma (EEG) y otros marcadores relevantes.

Uno de los principales atributos que distingue a los dispositivos fisiológicos es su capacidad para proporcionar mediciones objetivas de las respuestas emocionales y cognitivas. A diferencia de las evaluaciones basadas únicamente en autoinformes, que pueden estar sujetas a sesgos y variaciones individuales, los dispositivos fisiológicos ofrecen una ventana directa hacia las respuestas fisiológicas de la persona usuaria. Además, la naturaleza en tiempo real de las mediciones fisiológicas aporta un valor sustancial a la evaluación. Permiten capturar estas respuestas en el momento en que ocurren, proporcionando una comprensión detallada de los cambios emocionales y cognitivos que se generan en tiempo real.

La utilización de dispositivos fisiológicos no solo enriquece la recopilación de datos, sino que también abre la posibilidad de capturar respuestas sutiles y no conscientes. Las personas usuarias pueden no ser capaces de verbalizar con precisión sus estados emocionales y cognitivos durante la interacción con la tecnología. Los indicadores fisiológicos, como los cambios en la actividad eléctrica cerebral o la conductancia de la piel, pueden revelar aspectos internos que de otro modo permanecerían ocultos, lo que brinda una perspectiva más completa y holística de la experiencia de la persona usuaria.

En el contexto de HRI en entornos industriales, los dispositivos fisiológicos se presentan como una herramienta emergente para evaluar la UX y la aceptación de los robots por parte de las personas trabajadoras. Permiten examinar en tiempo real cómo la presencia y el comportamiento de los robots inciden en la experiencia subjetiva de las personas trabajadoras.

2.4.1.1 Dispositivos y métricas

A continuación, se presentan los dispositivos y métricas que se utilizan para la evaluación de las diversas respuestas fisiológicas y emocionales de las personas.

2.4.1.1.1 Actividad cardíaca

La actividad cardíaca refleja la reacción del sistema nervioso autónomo frente a estímulos físicos o emocionales externos o internos, tales como situaciones de estrés (Blandino, 2023). Por lo general, se mide mediante la técnica de electrocardiografía (ECG), la cual registra la actividad eléctrica del corazón a través de la aplicación de electrodos en configuraciones de posición estándar o mediante otros dispositivos portátiles como relojes inteligentes. Dentro de este contexto, el parámetro cardinal obtenido a través de la evaluación de la actividad cardíaca es la variabilidad de la frecuencia cardíaca (HRV, por sus siglas en inglés), un concepto que abarca tanto la frecuencia de latidos cardíacos por minuto (lat/min) como las oscilaciones en el intervalo temporal entre dos ondas R consecutivas en la señal de ECG (Castaldo et al., 2015). Además, es común la medición de otros indicadores tales como la

frecuencia cardíaca (HR, por sus siglas en inglés), así como el consumo individual de oxígeno (VO_2 , por sus siglas en inglés) y su punto máximo (VO_{2max}).

2.4.1.1.2 Actividad electrodérmica (EDA)

La actividad eléctrica de la piel, también identificada como respuesta galvánica de la piel (GSR, por sus siglas en inglés) o conductancia de la piel (SC, por sus siglas en inglés), refleja las cambiantes propiedades eléctricas de la dermis, siendo resultado de las oscilaciones en la producción de sudor del organismo humano. La señal EDA puede desglosarse en componentes de naturaleza tónica y fásica. El componente tónico, reconocido como nivel de conductancia de la piel (SCL, por sus siglas en inglés), alude a las variaciones de largo plazo en la señal EDA, las cuales se originan en las características fisiológicas inherentes al individuo. Por otro lado, el componente fásico de la señal, denominado respuesta de conductancia de la piel (SCR, por sus siglas en inglés), representa la reacción de la conductancia cutánea en respuesta a estímulos a corto plazo. Por esta razón, se considera un valioso indicador de los niveles de estrés experimentados (Setz et al., 2010).

2.4.1.1.3 Actividad respiratoria

La respiración se ve particularmente influenciada por el estado emocional y eventos de estrés, y ha sido reconocida como más sensible que otras medidas fisiológicas (Nicolò et al., 2020). Peruzzini et al. (2018, 2020) calcularon la frecuencia respiratoria, midiendo respiraciones por minuto (RPM) en combinación con la medición de la temperatura de la piel (ST), llevando a cabo un análisis completo de las condiciones de salud de la persona operaria en el entorno laboral con el propósito de evaluar el estrés y la fatiga.

2.4.1.1.4 Actividad cerebral

Usualmente, esta actividad se registra a través de la electroencefalografía (EEG). El EEG es una técnica adoptada para el registro de la actividad eléctrica generada por el cerebro. Los datos son recolectados mediante la aplicación de electrodos en el cuero cabelludo, generalmente según una disposición de posiciones estándar (Blandino, 2023). Cada electrodo mide el voltaje producido por la actividad neuronal de la región cerebral en la que se coloca. Bitbrain (2020) establece una conexión entre la actividad cerebral y diversas emociones, tales como la valencia emocional, la memorización, la atención y el grado de involucramiento (*engagement*).

2.4.1.1.5 Actividad muscular

La electromiografía (EMG) y los estudios de conducción nerviosa son pruebas que evalúan la actividad eléctrica presente en los músculos y los nervios. Estas pruebas analizan cómo los nervios emiten señales eléctricas para generar respuestas específicas en los músculos. Estas respuestas musculares, a su vez, generan señales detectables. En el caso de la electromiografía, se investigan las señales eléctricas emitidas por los músculos tanto cuando están en reposo como cuando están en acción. Por otro lado, los estudios de conducción nerviosa miden la velocidad y eficiencia con la que las señales eléctricas viajan a través de los nervios en el cuerpo. Aunque esta modalidad se utiliza principalmente en pruebas de dispositivos, ergonomía y factores humanos, si se conecta a los músculos faciales, también puede funcionar como otra forma de medir las emociones.

2.4.1.2 *Uso de los dispositivos fisiológicos en la literatura*

La incorporación de dispositivos fisiológicos en el ámbito académico y científico está adquiriendo un papel esencial en el análisis de la interacción humano-tecnología. Taffese (2017) llevó a cabo una revisión sistemática que abordó la UX, la fisiología, los estudios de las emociones, el sistema nervioso y las herramientas de medición correspondientes. El propósito

de su estudio fue dar una base sólida sobre el uso de herramientas fisiológicas en la investigación de la UX. El objetivo era ofrecer un punto de referencia donde se recopilase la comprensión de las implementaciones y estudios previos de herramientas fisiológicas, EEG y EMG en concreto. La exploración de Taffese (2017) abarcó un análisis profundo de los beneficios, limitaciones, contextos de aplicación, áreas de implementación y posibles desarrollos futuros asociados con estas tecnologías. Aunque estas herramientas proporcionan un medio para examinar las emociones en los tests de UX, el autor enfatizó la importancia de contextualizar los resultados considerando el entorno experimental, el contexto de uso y las aplicaciones o sistemas de *software* utilizados. Dado que las interpretaciones de las respuestas fisiológicas pueden ser ambiguas en ocasiones, Taffese recomendó el uso de enfoques combinados con otras herramientas para validar las inferencias emocionales.

En el estudio de Baig & Kavakli (2019), centraron su atención en señales fisiológicas para evaluar aspectos cognitivos y emocionales. Sus resultados mostraron una correlación significativa entre los datos de autorregistro y los datos fisiológicos, destacando la coherencia entre ambos conjuntos de información.

En la literatura, no se ha revisado cómo se aplican estas señales fisiológicas en los test de usuario a la hora de evaluar la UX. Es por ello por lo que realizamos una SLR con el objetivo de identificar testeos o casos de estudio de UX realizados con monitorización fisiológica hasta la actualidad (Apraiz et al., 2021).

De este modo, la SLR permitió identificar 8 artículos que utilizan medidas fisiológicas en soluciones app o web (Bai & Fuglerud, 2018; Barral et al., 2017; Foglia et al., 2014; Jiang et al., 2019; López-Gil et al., 2010; Mangion et al., 2020; Menges et al., 2019; Niksirat et al., 2019); 7 artículos en el ámbito del Gaming (Bian et al., 2016; Giakoumis et al., 2011; Holz et al., 2013; Huynh et al., 2018; Katsigiannis et al., 2019; Meza-Kubo et al., 2016; Stein et al., 2018); 3 artículos en Multimedia (Cabibihan & Chauhan, 2017; Gulliver & Ghinea, 2006; Y. C. Lee et al., 2016); 3 en realidad virtual (Gramann et al., 2010a; Slater et al., 2006a; Yan et al., 2017), 2 en *Brain Computer Interface* (Randolph & Jackson, 2010; Zickler et al., 2013) y 2 en industria (Peruzzini et al., 2018, 2020). También se identifican en campos como los artículos de ámbitos como el *Computer-Aided Design* (CAD) (Liu et al., 2014), *machine learning* (Alzoubi et al., 2012), *healthcare* (Sanches et al., 2019), *digital music instrument* (DMI) (Cincuegrani et al., 2016), sistemas colaborativos (Gupta et al., 2016), habilidades cognitivas (Nourbakhsh et al., 2017), *lifelogging* (Jiang et al., 2019) y políticas de privacidad (Karegar et al., 2020).

En el ámbito de la robótica, Harriot et al. (2013) miden la HR y la HRV. En el estudio de Gervasi et al. (2023) también se mide la HR y la EDA. Gervasi et al. (2022) emplea la pulsera Empática E4 biosensor de Empática (2022) para recolectar datos de EDA a 4 Hz, datos de frecuencia cardíaca (HR) a través de la PPG a 64 Hz y datos del acelerómetro de 3 ejes a 32 Hz. Grandi et al. (2022) proponen el uso de la medición de HR, HRV y PD. Khamaisi, Grandi et al. (2022) además de estas últimas, proponen también la medición de EDA. Akalin et al. (2022) utilizan un análisis de expresión facial mediante Affdex y la pulsera Empática E4 también.

Los dispositivos fisiológicos utilizados en los estudios identificados fueron:

- El EEG se presenta como el dispositivo más utilizado, ya que se utiliza en los siguientes estudios: (Bai & Fuglerud, 2018; Barral et al., 2017; Cincuegrani et al., 2016; Gramann et al., 2010; Holz et al., 2013; Jiang et al., 2019; Y. C. Lee et al., 2016; Liu et al., 2014; Mangion et al., 2020; Menges et al., 2019; Meza-Kubo et al., 2016; Niksirat et al., 2019; Stein et al., 2018; Yan et al., 2017; Zickler et al., 2013). Sin embargo, no se identifica ningún estudio que implemente el EEG en la HRI.
- La segunda herramienta más utilizada es el GSR, y se utiliza en los siguientes estudios: (Alzoubi et al., 2012; Barral et al., 2017; Cabibihan & Chauhan, 2017; Foglia et al., 2014; Giakoumis et al., 2011; Huynh et al., 2018; Jiang et al., 2019; Katsigiannis et al.,

2019; Liu et al., 2014; Menges et al., 2019; Nourbakhsh et al., 2017; Randolph & Jackson, 2010; Slater et al., 2006a).

- Se identificó el uso de eye tracker en los siguientes artículos: (Bai & Fuglerud, 2018; Gulliver & Ghinea, 2006; Gupta et al., 2016; Karegar et al., 2020; Y. C. Lee et al., 2016; López-Gil et al., 2010; Menges et al., 2019; Nourbakhsh et al., 2017; Peruzzini et al., 2018, 2020).
- Se ha identificado el uso de ECG en los siguientes estudios: (Alzoubi et al., 2012; Barral et al., 2017; Bian et al., 2016; Cincuegrani et al., 2016; Giakoumis et al., 2011; Katsigiannis et al., 2019; Nourbakhsh et al., 2017; Slater et al., 2006a). Además, se ha identificado también el uso de ECG en estudios relacionados con la HRI como Gervasi et al. (2023), Grandi et al. (2022), Khamaisi, Grandi, et al. (2022) y Akalin et al. (2022).
- Se ha identificado el uso de EMG en los siguientes estudios: (Alzoubi et al., 2012; Bian et al., 2016; Liu et al., 2014).

Con el propósito de discernir las herramientas que se emplean en conjunción en las evaluaciones de UX. Se ha procedido a efectuar un registro sistemático de la frecuencia de co-utilización de las mismas. Dentro de este contexto, emerge que el binomio de herramientas más preponderante consiste en la ECG en combinación con la GSR, el cual ha sido aplicado de manera simultánea en un total de siete instancias. En una posición subsiguiente, se constata la conjunción del ECG y la EEG, la cual ha sido empleada en tres ocasiones. En un ámbito análogo, se identifica la co-aplicación del *Eyetracker* y el EEG en dos casos, así como la confluencia entre la GSR y la EEG, así como entre la EMG y la GSR, ambas materializadas en dos escenarios. Finalmente, también se constata la simultaneidad en la utilización de la ECG y la EMG en dos ocasiones.

Con lo que respecta a la muestra de personas usuarias utilizada en los casos de estudio identificados en la literatura, la media es de 24 personas y la desviación típica de 22,6. Lo que significa que el tamaño de la muestra de personas usuarias varía según el caso de estudio. El testeo realizado con menor cantidad de personas tiene un tamaño de 4 personas y son los realizados por Holz et al. (2013), Zickler et al. (2013) y Peruzzini et al. (2020a). Al contrario, la muestra más amplia utilizada es de 108 personas realizado por Gulliver & Ghinea (2006).

Se ha realizado un cálculo para conocer cuál es el promedio de participantes en los casos de estudio identificados en la literatura dependiendo de la herramienta fisiológica utilizada. Tal y como se puede ver en la Figura 23, se han obtenido los siguientes resultados:

- En el caso del EEG, la mediana de la muestra es de 18 participantes y el valor medio de 19,94. El 50% de los casos de estudio realizados con EEG tienen una muestra de 9,5 hasta 24,5 participantes.
- En el caso del GSR/EDA, la mediana de la muestra es de 25 participantes y el valor medio de 26,31. El 50% de los casos de estudio realizados con GSR/EDA tienen una muestra desde 18,5 hasta 32,5.
- Respecto al *Eyetracker*, la mediana de la muestra de personas es de 18 y el valor medio de 32,09. El 50% de los casos de estudio realizados con *Eyetracker* tienen una muestra desde 12 hasta 54.
- En cuanto al ECG, la mediana de la muestra de personas usuarias es de 25 y el valor medio de 28,89. El 50% de los casos de estudio realizados con ECG tienen una muestra de 19,5 hasta 35,5.
- Finalmente, en el caso del EMG al solo identificarse 3 casos de estudio, se identifica que la mediana es el caso de estudio que tiene una muestra de 27 personas, el límite inferior el que tiene 15 y el superior el de 36. El valor medio es de 26.

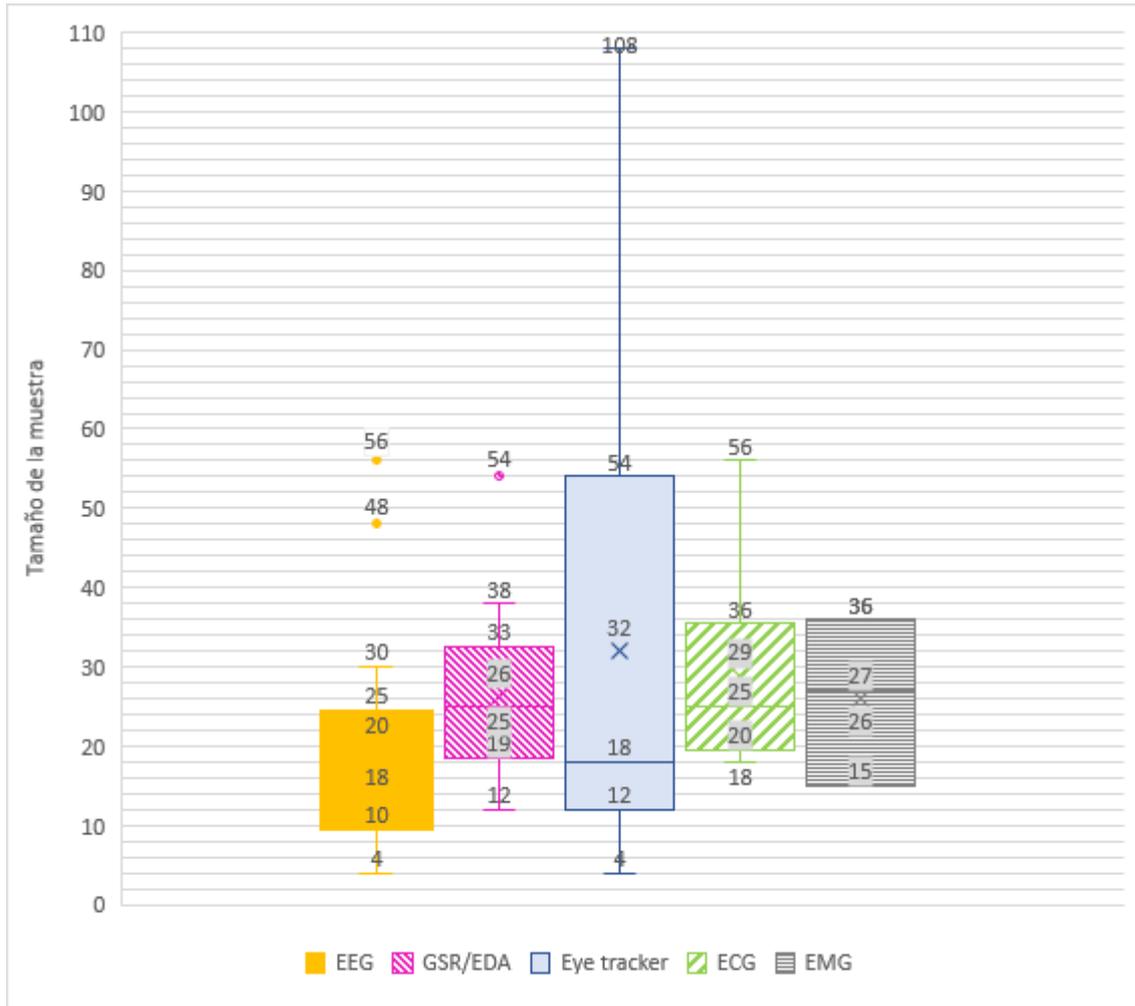


Figura 23: Tamaño de la muestra según la herramienta fisiológica.

2.4.2 Evaluación de la percepción de usuario

Los cuestionarios representan una herramienta de uso extendido en la evaluación de la percepción de la persona usuaria, así como en la calidad y facilidad de uso de programas informáticos (Laugwitz et al., 2008). Permiten una medición cuantitativa eficiente de las características del producto ya que se miden habitualmente mediante escalas de Likert o pares semánticos. Resultan particularmente eficaces para la aplicación y análisis, cumpliendo con los requisitos de evaluación previamente abordados en la sección 1.2.4, conforme a lo propuesto por Väänänen-Vainio-Mattila et al. (2008).

No obstante, es importante destacar que la aplicación de algunos cuestionarios puede llevar bastante tiempo, como es el caso de PANAS-X (Watson & Clark, 1999) o PUTQ (Lin et al., 1997), que están constituidos por 60 y 100 ítems respectivamente. Se utilizan para medir la perspectiva de usuario y no requieren necesariamente de ningún tipo de monitorización. La evaluación de la UX con cuestionarios es de interés tanto para los investigadores que realizan trabajos y estudios académicos como para los profesionales de la UX que trabajan en la industria.

2.4.2.1 Cuestionarios estandarizados

En el marco de esta investigación, se realizó un Mapeo Sistemático (SM, por sus siglas en inglés) con el propósito de identificar cuestionarios relacionados con la UX (Apraiz & Lasa, 2020). La revisión nos permitió la identificación de un total de 32 cuestionarios estandarizados, abarcando el período comprendido entre 1988 y 2020. Además de catalogarlos, se exploró la naturaleza de estos instrumentos, su proceso de desarrollo, integración en la comunidad académica y su empleo a lo largo del tiempo.

Los cuestionarios identificados fueron: *Positive and Negative Affect Schedule* (PANAS) (Watson et al., 1988), *Questionnaire for User Interface Satisfaction* (QUIS) (Chin et al., 1988), *Perceived Usefulness and Ease of Use* (Davis, 1989), *Printer Scenario Questionnaire* (PSQ) (J. R. Lewis, 1991a), *After-Scenario Questionnaire* (ASQ) (J. R. Lewis, 1991b), *Post-Study System Usability Questionnaire* (PSSUQ) (J. R. Lewis, 1992b), *Computer System Usability Questionnaire* (CSUQ) (J. R. Lewis, 1992a), *Software Usability Measurement Inventory* (SUMI) (Kirakowski & Corbett, 1993), *Positive and Negative Affect Schedule – X* (PANAS-X) (Watson & Clark, 1999), *System Usability Scale* (SUS) (Brooke, 1996), *Purdue Usability Testing Questionnaire* (PUTQ) (Lin et al., 1997), *Website Analysis Measurement Inventory* (WAMMI) (Kirakowski & Cierlik, 1998), Cuestionario USE (Lund, 2001), *AttrakDiff* (Hassenzahl, 2001), *Web Quality* (WQ) (Aladwani & Palvia, 2002), *AttrakDiff 2* (Hassenzahl et al., 2003), *Emotions in Consumer Behaviour* (Laros & Steenkamp, 2005), *Website Usability* (WU) (J. Wang & Senecal, 2007), *User Experience Questionnaire* (UEQ) (Laugwitz et al., 2008), *Intranet Satisfaction Questionnaire* (ISQ) (Bargas-Avila et al., 2009), *Usability Metric for User Experience* (UMUX) (Finstad, 2010), *Visual Aesthetics of Websites Inventory* (VisAWI) (Moshagen & Thielsch, 2010), *DEEP: Design-oriented evaluation of perceived usability* (T. Yang et al., 2012), *Visual Aesthetics of Websites Inventory Short* (VisAWI-S) (Moshagen & Thielsch, 2013), *Usability Metric for User Experience LITE* (UMUX LITE) (J. R. Lewis et al., 2013), *Emotional Metrics Outcome* (EMO) (J. R. Lewis & Mayes, 2014), *Alternate Usability* (AltUsability) (J. R. Lewis et al., 2015), *Standardized User Experience Percentile Rank Questionnaire* (SUPR-Q) (Sauro, 2015), *Standardized User Experience Percentile Rank Questionnaire mobile* (SUPR-Qm) (Sauro & Zarolia, 2017), *User Experience Questionnaire Short* (UEQ-S) (Schrepp et al., 2017), *modular evaluation of key Components of User Experience* (meCUE) (Minge et al., 2017), y *User Experience Questionnaire +* (UEQ+) (Schrepp & Thomaschewski, 2019).

Como resultado de este análisis, se identificaron los cinco cuestionarios más frecuentemente utilizados en la evaluación de la UX:

1. **Perceived Usefulness and Ease of Use** (Davis, 1989): Este cuestionario mide la usabilidad y los siguientes sub-constructos: utilidad y facilidad de uso.
2. **Positive and Negative Affect Schedule (PANAS)** (Watson et al., 1988): Este cuestionario se centra en evaluar las emociones dividiéndolas en afecto positivo y afecto negativo.
3. **System Usability Scale (SUS)** (Brooke, 1996): Mide la usabilidad y la aprendibilidad.
4. **Positive and Negative Affect Schedule – X (PANAS-X)** (Watson & Clark, 1999): Este cuestionario se concentra en la medición de las emociones desglosándolos en *Afecto positivo, afecto negativo, escalas básicas de emociones negativas, escalas básicas de emociones positivas y otras escalas afectivas*.
5. **Questionnaire for User Interface Satisfaction (QUIS)** (Chin et al., 1988): Se enfoca en la **usabilidad** y comprende sub-constructos como *factores de la pantalla, aprendibilidad y capacidad del sistema*.

Es importante señalar que estos cuestionarios, a pesar de su amplia utilización, tienden a medir componentes específicos, como usabilidad o emociones, y no abordan la experiencia en su totalidad. Esta limitación podría atribuirse, en parte, a la longevidad de los cinco cuestionarios más empleados, ya que se encuentran entre los más antiguos de los 32 identificados. Además, la complejidad de evaluar la UX de manera holística, como sugieren Luther et al. (2020), también podría influir en esta tendencia, llevando a los profesionales a centrarse en aspectos particulares de la UX, en lugar de considerarla en su totalidad.

Por otra parte, en lo referente a los constructos y sub-constructos que miden cada uno de los cuestionarios, existen diferentes. Por ejemplo, el cuestionario datado como más antiguo, el QUIS (Chin et al., 1988), evalúa la usabilidad y cuatro sub-constructos, mientras que el más reciente, UEQ+ (Schrepp & Thomaschewski, 2019) abarca siete sub-constructos solo en usabilidad. La tendencia general es que los cuestionarios más recientes tienden a medir un mayor número de sub-constructos, a excepción de las versiones abreviadas, como UMUX LITE (J. R. Lewis et al., 2013) o UEQ-S (Schrepp et al., 2017). Sin embargo, se siguen manteniendo los sub-constructos iniciales, con lo que se puede concluir, que estos se han desglosado y que se han ido añadiendo nuevos a medida que el campo de la UX se ha ido desarrollando y estudiando.

Entre los cuestionarios identificados como más completos, que evalúan la UX de manera integral, se encuentran: AttrakDiff 2 (Hassenzahl et al. (Sauro & Zorolia, 2017) o, 2015), SUPR-Qm (Sauro & Zorolia, 2017) y UEQ+ (Schrepp & Thomaschewski, 2019).

En cuanto a las e(Sauro & Zorolia, 2017) o, 2015), SUPR-Qm (Sauro & Zorolia, 2017) y UEQ+ (Schrepp & Thomaschewski, 2019) miden la confianza. AttrakDiff 2 (Hassenzahl et al., 2003), mide tanto la estimulación como la identidad de la calidad hedónica, conceptos que, aunque únicamente se han identificado en este cuestionario, podrían ser equiparables a la estimulación o confianza presentes en UEQ+ (Schrepp & Thomaschewski, 2019).

En un ámbito más relacionado a la usabilidad, AttrakDiff 2 (Hassenzahl et al., 2003) evalúa la calidad pragmática, un concepto que, al igual que en el caso de las emociones, solo se emplea en este cuestionario, pero que podría ser similar al concepto de usabilidad que lo miden de forma genérica tanto SUPR-Q (Sauro, 2015) como SUPR-Qm (Sauro & Zorolia, 2017). Sin embargo, UEQ+ (Schrepp & Thomaschewski, 2019) desglosa el concepto en sub-constructos como aprendibilidad, utilidad, calidad de contenido, eficiencia, controlabilidad, uso intuitivo y adaptabilidad.

En lo referente a la apariencia, el único constructo comúnmente medido en los cuatro cuestionarios es la atraktividad, aunque UEQ+ (Schrepp & Thomaschewski, 2019) también evalúa la estética y el valor. UEQ+ (Schrepp & Thomaschewski, 2019) destaca como el cuestionario más completo y reciente, ya que mide mayor cantidad de sub-constructos. Dado que UEQ+ es modular, permite que los profesionales personalicen el cuestionario según sus necesidades específicas, adaptándose de manera más efectiva a diversas situaciones.

Uno de los requisitos fundamentales para una evaluación efectiva de la UX es la rapidez, la ligereza y la eficiencia del método (Väänänen-Vainio-Mattila et al., 2008). En este contexto, se analizó el número de ítems de cada cuestionario, reconociendo que, a mayor cantidad de ítems, el cuestionario se vuelve más tedioso. El cuestionario con mayor número de ítems es PUTQ (Lin et al., 1997), compuesto de 100 ítems. Seguido, está el cuestionario UEQ+ (Schrepp & Thomaschewski, 2019) con 64 ítems. Sin embargo, es importante destacar que UEQ+ es modular y consta de 16 constructos, cada uno de ellos con 4 ítems, permitiendo que el profesional seleccione los sub-constructos según sus necesidades específicas. Por lo tanto, los 64 ítems representarían el máximo en caso de abordar todos los constructos.

PANAS-X (Watson & Clark, 1999) consta de 60 pares semánticos, mientras que SUMI (Kirakowski & Corbett, 1993) se compone de 50 ítems. Al contrario, UMUX LITE (J. R. Lewis et al., 2013) es el cuestionario más corto, ya que se compone de 2 ítems. Seguido de él, PSQ (J. R. Lewis, 1991a) y ASQ (J. R. Lewis, 1991b) se componen de 3 ítems y VisAWI-S (Moshagen & Thielsch, 2013) de 4 ítems.

Se ha podido observar que la tendencia es que los cuestionarios están compuestos de cada vez menos ítems, y que en los últimos años se han creado versiones reducidas de cuestionarios existentes, como pueden ser UEQ-S (Schrepp et al., 2017), VisAWI-S (Moshagen & Thielsch, 2013) o UMUX-LITE (J. R. Lewis et al., 2013). Esta tendencia puede ser debido a que algunos cuestionarios resultan demasiado extensos para algunos contextos; por ejemplo, los que no requieran de información detallada, o las investigaciones en las que la persona usuaria debe rellenar más de una vez el mismo cuestionario para comparar los resultados de diferentes objetos de estudio. Estas versiones cortas suelen estar compuestas desde 2 hasta 8 ítems, menor número que en sus versiones originales, y, en consecuencia, el tiempo que las personas usuarias necesitan para cumplimentarlo también es menor. Sin embargo, las versiones cortas también tienen su desventaja, ya que se pierden algunos aspectos detallados que se incluyen en su versión original.

UEQ (Laugwitz et al., 2008) en su versión original, da una respuesta detallada sobre la *aprendibilidad, eficiencia, controlabilidad, atraktividad y originalidad*. Esto se pierde en su versión corta que solo distingue entre la *calidad pragmática* y la *calidad hedónica (estimulación e identidad)*. A pesar de este hecho, los autores demostraron que la versión corta es capaz de predecir el comportamiento de la versión completa en cuanto a la *calidad pragmática y hedónica* (Schrepp et al., 2017). Sin embargo, los autores también incluyen que UEQ-S está pensada para escenarios específicos que no permiten el empleo de un completo UEQ, y que no permite medir las detalladas cualidades.

En el caso de VisAWI-S (Moshagen & Thielsch, 2013), los resultados indican que la versión corta es una medida fiable y que proporciona una buena aproximación a la versión completa. Sin embargo, también se pierde la información detallada que ofrece la versión original. Esto hace que resulte preferible utilizar la versión original cuando se requiere de información detallada.

Por lo tanto, puede concluirse que las versiones cortas sólo son válidas para ciertos contextos, y que son los profesionales de la UX los que deberán valorar si el equilibrio entre el tiempo requerido y la información obtenida es adecuado.

2.4.2.2 Cuestionarios utilizados en estudios de HRI

En adición a los cuestionarios específicos de UX previamente mencionados, es importante destacar que los estudios experimentales centrados en la HRI hacen uso de una variedad de cuestionarios destinados a medir distintos aspectos de esta interacción. Estos instrumentos se han identificado en la SLR y se describen a continuación:

- **Cuestionario Godspeed** (Bartneck et al., 2008): Este cuestionario se utiliza con el propósito de medir múltiples facetas de la percepción de robots, incluyendo el antropomorfismo, la animación, la simpatía, la inteligencia percibida y la seguridad percibida. Se encontró que este cuestionario fue empleado en dos ocasiones específicas, en los estudios realizados por Schillaci et al. (2013) y Joosse et al. (2021).
- **Cuestionario SUS** (System Usability Scale) (Brooke, 1996): El Cuestionario SUS se utiliza para medir la usabilidad y la simpatía de la interacción con robots. Fue empleado en el estudio de Danielsson et al. (2017).
- **Cuestionario basado en el Cuestionario de Satisfacción de Usabilidad de Computadoras IBM**: En el estudio de Almeida et al. (2020), se utilizó un cuestionario basado en el Cuestionario de Satisfacción de Usabilidad de Computadoras IBM para medir la usabilidad y la satisfacción en la interacción con robots.
- **Cuestionario Autogenerado**: En el estudio de Lasota & Shah (2015), se diseñó un cuestionario autogenerado para evaluar la satisfacción de las personas participantes con los robots como compañeros de equipo, así como para determinar la percepción de seguridad y comodidad.
- **SE-HRI** (von der Pütten & Bock, 2018): Está diseñado específicamente para evaluar la percepción de autoeficacia de las personas usuarias cuando interactúan con un robot.

En estudios más recientes, se han empleado otros cuestionarios para evaluar aspectos específicos de la interacción HRI. En una investigación llevada a cabo por Gervasi et al. (2022), se recurrió al empleo del cuestionario NARS (Nomura et al., 2004) para evaluar la actitud de las personas participantes hacia los robots, y se utilizó la Self-Assessment Manikin (SAM) para cuantificar las respuestas afectivas. En otro estudio conducido por Gervasi et al. (2023), se optó por utilizar el cuestionario NASA-TLX y nuevamente el SAM. Estos instrumentos se emplearon con el propósito de recopilar información relativa a las dimensiones de demanda mental, demanda física y demanda temporal, así como para evaluar el estado afectivo autoinformado de las personas participantes en el contexto de la HRI. Además, se utilizó un cuestionario específico sobre la calidad de la interacción para evaluar la percepción de las personas participantes en cuanto a diversos aspectos de la interacción con el cobot. En el estudio de Grandi et al. (2022), en el que se propone la creación del UXI, proponen también el uso del cuestionario NASA-TLX. En el estudio de Beschi et al. (2020) utilizan un cuestionario que evalúa el nivel de riesgo e interferencia percibido por la persona.

Adicionalmente, en el estudio de Akalin et al. (2022), se empleó el Big Five Inventory (Rammstedt & John, 2007) al inicio del experimento para determinar la personalidad de las personas usuarias. Entre los cuestionarios utilizados en diferentes momentos del experimento se incluyen un cuestionario sobre el confort basado en el estudio de Kim & Mutlu (2014), un cuestionario sobre la seguridad percibida (Akalin et al., 2019), un cuestionario sobre el sentido de control adaptado de Strube & Werner (1984), y el cuestionario Trust Perception Scale-HRI (Schaefer, 2016). El cuestionario post-ejecución fue el de Weiss et al. (2009).

2.5 Estudio crítico del estado del arte

En esta sección, se llevará a cabo una revisión crítica del estado del arte en el campo de estudio. Se abordarán diferentes aspectos, comenzando por el examen de los factores humanos más influyentes en la adopción de robots en contextos industriales. A continuación, se analizarán los métodos utilizados para evaluar estos factores humanos, incluyendo una comparación de estudios experimentales y una discusión sobre las herramientas y técnicas de evaluación identificadas, con el propósito de analizar su aplicabilidad y eficacia en la evaluación de la HRI en entornos industriales. Por último, se reflexionará sobre la implicación de las variables cinemáticas y las interfaces en la percepción de las personas.

2.5.1 Factores humanos influyentes en la aceptación de los robots por parte de las personas en los contextos industriales

Para comprender y mejorar la relación entre las personas y los robots en los entornos industriales, es fundamental conocer qué factores humanos influyen en la aceptación de los robots por parte de las personas.

2.5.1.1 *Percepción de riesgo y seguridad en la aceptación de la HRI en entornos industriales*

La literatura en el campo de la seguridad en la HRI se centra en garantizar la seguridad de las personas operarias al prevenir y controlar situaciones de riesgo potenciales en las interacciones entre personas y robots (Gualtieri et al., 2021). Sin embargo, es importante destacar que, más allá de esta perspectiva centrada en el robot, existe un aspecto igualmente crítico: la seguridad percibida. Esta se refiere a la evaluación subjetiva que una persona realiza sobre la seguridad de una tecnología o sistema.

En este contexto, es importante resaltar que dos de los factores más influyentes en la aceptación de los sistemas robóticos son el riesgo y los beneficios percibidos (Meissner et al., 2020). Estos factores se traducen en el entorno industrial en la percepción de seguridad y utilidad de la tecnología robótica. La seguridad, junto con la ergonomía, constituye uno de los aspectos más ampliamente investigados en el campo de la HRI (Gualtieri et al., 2021).

En el estudio experimental realizado por Beschi et al. (2020), se evalúa el riesgo percibido a través de un cuestionario diseñado específicamente por los propios autores. La percepción de riesgo juega un papel crucial en cómo las personas usuarias aceptan y confían en los sistemas robóticos en entornos industriales. Asimismo, su homónimo, la seguridad percibida se ha abordado en estudios como los de Schillaci et al. (2013) y Joosse et al. (2021), donde se utilizaron el cuestionario estandarizado Godspeed para evaluar esta dimensión, o el estudio de Lasota & Shah (2015), que utilizaron un cuestionario autogenerado. La consideración y medición de la percepción de riesgo y seguridad percibida son esenciales para comprender cómo las personas perciben y responden a la presencia de robots en el contexto industrial. Estos factores influyen en gran medida en la aceptación y la actitud hacia la tecnología robótica, y, por lo tanto, tienen un impacto directo en la eficacia y la eficiencia de la interacción humano-robot en entornos de trabajo.

En el contexto industrial, la seguridad percibida implica la percepción de las personas trabajadoras de que interactuar con un robot o sistema robótico no conlleva un riesgo significativo para su integridad física. Esta percepción subjetiva de seguridad desempeña un papel crucial en la aceptación de la tecnología, ya que las personas tienden a mostrar resistencia o rechazo hacia sistemas que perciben como peligrosos o inseguros, como han señalado Meissner et al. (2020).

Aunque la seguridad en las interacciones entre personas y robots es una preocupación fundamental, es igualmente importante considerar cómo las personas perciben y evalúan esta

seguridad. Además, la percepción de riesgo y seguridad percibida influyen directamente en la actitud y la disposición de las personas trabajadoras hacia la tecnología robótica, lo que a su vez tiene un impacto en la eficacia y la eficiencia de la HRI en entornos de trabajo industriales.

En este contexto, surge la oportunidad de avanzar en la investigación y desarrollo de métodos más precisos y sensibles destinados a evaluar la seguridad percibida en entornos industriales. Esto implica la posible creación de escalas y cuestionarios específicos diseñados para capturar las dimensiones esenciales relacionadas con la seguridad percibida, permitiendo así una evaluación más minuciosa y detallada. Además, es crucial explorar las evaluaciones que actúan como moderadoras en la relación entre la percepción de riesgo y la aceptación de la tecnología robótica. Estas evaluaciones pueden abarcar factores individuales, contextuales y culturales que influyen en cómo las personas evalúan la seguridad de los robots y, en consecuencia, su disposición para utilizarlos.

Por otro lado, se plantea la cuestión de determinar cuáles son las estrategias más efectivas para comunicar la seguridad de los robots a las personas trabajadoras, lo que podría tener un impacto significativo en su seguridad percibida. Este enfoque implica la investigación y evaluación de prácticas y métodos de comunicación específicos que puedan mejorar la percepción de seguridad de las personas trabajadoras en relación con la tecnología robótica.

2.5.1.2 Ergonomía cognitiva y bienestar en la HRI

La ergonomía, un factor ampliamente estudiado en el campo de la HRI, desempeña un papel fundamental en la configuración de entornos de trabajo seguros y eficientes (Gualtieri et al., 2021). Los estudios sobre ergonomía en la HRI a menudo se dividen en dos categorías principales, "Ergonomía física" y "Ergonomía cognitiva y organizacional", ambas diseñadas para abordar aspectos cruciales relacionados con el bienestar y la eficiencia de las personas en entornos industriales automatizados.

La ergonomía física se enfoca en la reducción de la carga biomecánica que experimentan las personas trabajadoras durante sus interacciones con sistemas robóticos. Esto implica la consideración de aspectos como la postura corporal, el esfuerzo físico requerido para realizar tareas, la prevención de lesiones musculoesqueléticas y la optimización de la biomecánica. En esencia, la ergonomía física busca garantizar que las tareas que implican el trabajo conjunto entre humanos y robots se realicen de manera segura y sin imponer un exceso de estrés físico a las personas operarias.

A su vez, la ergonomía cognitiva se centra en el bienestar mental y la eficiencia en el entorno de trabajo. Aborda la forma en que se diseñan las interfaces y las interacciones para minimizar la carga cognitiva de las personas usuarias. Esto incluye aspectos como la claridad de la información presentada, la facilidad de comprensión de los sistemas y la reducción de la fatiga mental asociada a la interacción con tecnologías complejas. Además, la ergonomía organizacional se enfoca en la optimización de los sistemas organizativos y políticos para garantizar que la colaboración entre humanos y robots sea fluida y eficaz.

Si bien existen diversos estudios que buscan mejorar la ergonomía física en los puestos de trabajo, este trabajo se centrará más en la ergonomía cognitiva y el bienestar de las personas operarias. Estos aspectos son fundamentales para garantizar que la interacción entre personas y robots en entornos industriales sea no solo segura desde el punto de vista físico, sino también eficiente y cómoda desde el punto de vista cognitivo y emocional de las personas trabajadoras.

Es importante destacar que la ergonomía cognitiva también está estrechamente relacionada con la facilidad de uso de un sistema. Cuando un sistema robótico es fácil de entender y utilizar, las personas trabajadoras pueden interactuar con él de manera más eficiente, lo que a su vez influye en la percepción de la experiencia del usuario. Además,

reducir la carga cognitiva y facilitar la comprensión de las tareas contribuye a minimizar los errores humanos y a mejorar la seguridad en el lugar de trabajo.

La carga mental, la memoria y la valencia emocional son dimensiones a considerar en la medición de la ergonomía cognitiva. La carga mental se refiere a la cantidad de esfuerzo cognitivo requerido para llevar a cabo una tarea. Una alta carga mental puede llevar a la fatiga mental y afectar negativamente el bienestar de las personas. La memoria, por su parte, influye en cómo las personas retienen y utilizan la información en su trabajo con sistemas robóticos. La valencia emocional, que está relacionada con la carga emocional, también desempeña un papel importante en el bienestar. Las emociones positivas o negativas experimentadas durante la interacción con robots pueden influir en la percepción general de bienestar.

Cabe mencionar que la carga mental, la memoria y la valencia emocional pueden medirse y evaluarse mediante técnicas como el EEG, que proporciona insights valiosos sobre el estado cognitivo y emocional de las personas usuarias durante la HRI. Comprender estas dimensiones y su relación con la ergonomía cognitiva es esencial para diseñar sistemas que no solo sean seguros desde el punto de vista físico, sino que también promuevan el bienestar mental y emocional de las personas trabajadoras en entornos industriales automatizados. Esto plantea una interesante oportunidad de investigación, ya que, a pesar de que el EEG se ha consolidado como el dispositivo más empleado en estudios de UX con monitorización psicofisiológica, hasta el momento no ha sido utilizado en ningún estudio experimental que explore la interacción con robots.

2.5.1.3 La influencia de la utilidad percibida y facilidad de uso percibida en la actitud e intención de uso en la HRI en entornos industriales

La utilidad percibida se refiere a la evaluación subjetiva que una persona realiza sobre la eficacia y los beneficios de un sistema o tecnología. En el contexto de la HRI industrial, esta percepción se relaciona con la idea de que el robot aporta un valor tangible a las tareas laborales o mejora la eficiencia en el entorno de trabajo. La utilidad desempeña un papel esencial en la aceptación de un robot, ya que influye significativamente en la actitud de las personas y en su intención de utilizar la tecnología. Por lo tanto, la percepción de que un robot es efectivo y beneficioso no solo afecta a la disposición de las personas para utilizarlo, sino que también influye en gran medida en su actitud hacia la tecnología y en su propensión a adoptarla de manera efectiva en el entorno laboral.

La facilidad de uso percibida se refiere a la evaluación subjetiva que una persona realiza sobre cuánto de sencillo o conveniente percibe que es interactuar y utilizar una tecnología o sistema en particular. Es una medida de la percepción de un individuo acerca de la facilidad con la que puede aprender a utilizar una tecnología, realizar tareas con ella y resolver posibles problemas que puedan surgir durante su uso. En el contexto de la HRI en entornos industriales, la facilidad de uso percibida se relaciona con el grado de intrusividad y complejidad percibidas por una persona trabajadora al interactuar con un robot. Una alta facilidad de uso percibida puede aumentar la disposición de las personas a colaborar con robots en su entorno laboral, ya que se sienten cómodas y confiadas en utilizar esta tecnología en sus tareas diarias.

El modelo HRCAM (Bröhl et al., 2019), destaca específicamente la importancia de la utilidad percibida en la aceptación de los robots en entornos industriales. Según este modelo, el factor más influyente en la percepción de utilidad es la relevancia del trabajo, seguido de la norma subjetiva. Esta relevancia laboral puede atribuirse al contexto laboral e industrial en el que se implementan los robots. Las personas trabajadoras tienden a valorar más a un robot cuando perciben que este contribuye de manera significativa a su trabajo y mejora su eficiencia.

Además, la facilidad de uso percibida, otro factor crucial para la aceptación de la tecnología se ve influenciada por varias variables. Entre ellas se encuentran la autoeficacia, la afinidad con la tecnología, la percepción de control externo y la edad. Estos factores tienen un impacto en cómo las personas perciben la facilidad con la que pueden interactuar con los robots. Una mayor autoeficacia y una mayor afinidad con la tecnología suelen estar asociadas con una mayor percepción de facilidad de uso.

La adopción efectiva de robots en entornos industriales se encuentra intrínsecamente ligada a la percepción que las personas trabajadoras tienen acerca de la *utilidad y facilidad de uso* de estos dispositivos (Bröhl et al., 2019). Estas percepciones ejercen un impacto directo sobre la actitud de las personas trabajadoras y su disposición a emplear robots en su entorno laboral. En esta línea, la intención de uso se relaciona con la disposición de una persona a utilizar un sistema o tecnología en el futuro. Ambos conceptos están estrechamente ligados, ya que una actitud positiva generalmente conduce a una mayor intención de usar la tecnología. En el contexto de la HRI, una actitud favorable y una alta intención de uso son indicativos de la aceptación de los robots en el entorno laboral. Para diseñar sistemas robóticos que se ajusten a las necesidades y expectativas de las personas usuarias en contextos industriales, se vuelve fundamental una comprensión profunda de estos factores.

Hasta la fecha, la evaluación de la utilidad y la facilidad de uso se ha realizado principalmente a través de cuestionarios como el USE o el UEQ+, entre otros (Apraiz & Lasa, 2020). Sin embargo, no existe una herramienta específica diseñada para evaluar estos constructos en el contexto de la HRI en entornos industriales. Esto representa una oportunidad para desarrollar nuevas metodologías y herramientas que aborden de manera precisa y eficiente las percepciones de los trabajadores sobre la utilidad y la facilidad de uso de los robots en su vida laboral, impulsando así la evolución y la efectiva implementación de la robótica en la industria.

2.5.1.4 El impacto de aprendibilidad y controlabilidad en la aceptación en la HRI

La aprendibilidad es esencial en la HRI, ya que los robots y los sistemas de automatización a menudo requieren una interacción fluida y efectiva desde el primer encuentro. Las personas usuaria, que pueden ser personas trabajadoras en entornos industriales, deben poder aprender rápidamente cómo operar y comunicarse con estos robots. Según Joyce (2019), la aprendibilidad aborda la facilidad con la que las personas usuarias pueden llevar a cabo una tarea al interactuar por primera vez con una interfaz, así como la cantidad de repeticiones necesarias para lograr eficiencia en esa tarea. Un bajo nivel de aprendibilidad puede llevar a la frustración y la resistencia por parte de las personas usuarias, lo que a su vez podría obstaculizar la adopción exitosa de la tecnología.

La observación de cómo las personas usuarias perciben la aprendibilidad a lo largo del tiempo es esencial. En un contexto de HRI industrial, es común que las personas trabajadoras experimenten una curva de aprendizaje inicial al interactuar con tecnologías robóticas. Sin embargo, su percepción de la capacidad de aprendizaje puede evolucionar a medida que se familiarizan más con la tecnología y adquieren experiencia. Sin embargo, en la literatura existente no se han identificado estudios longitudinales que evalúen esta evolución de manera detallada en la HRI industrial. La realización de este tipo de estudio podría ofrecer información valiosa para diseñar sistemas que se ajusten de manera más efectiva a las cambiantes necesidades de las personas usuarias en el entorno industrial.

La aprendibilidad se ha explorado en estudios como el realizado por Danielsson et al. (2017), donde se utilizó el cuestionario SUS para su medición. Además del cuestionario SUS, existen múltiples cuestionarios que lo miden, tal y como se recoge en el mapeo sistemático de Apraiz & Lasa (2020). Estas incluyen QUIS (Chin et al., 1988), SUMI (Kirakowski & Corbett, 1993), PUTQ (Lin et al., 1997), WAMMI (Kirakowski & Cierlik, 1998), USE (Lund, 2001), UEQ

(Laugwitz et al., 2008), AltUsability (J. R. Lewis et al., 2015) y UEQ+ (Schrepp & Thomaschewski, 2019).

Otro de los factores a considerar, la controlabilidad, se refiere al grado en que una persona se siente en control de una tecnología y sus acciones, así como su capacidad para ajustar su comportamiento de acuerdo con sus necesidades o preferencias (Sarker y Wells, 2003). Se trata de un aspecto crítico en la HRI, ya que las personas usuarias necesitan sentir que tienen control sobre la tecnología con la que interactúan, especialmente en entornos laborales. La capacidad de ajustar la tecnología de acuerdo con las necesidades individuales o las preferencias de las personas usuarias es esencial para garantizar una colaboración efectiva entre personas y robots.

La percepción de controlabilidad también puede cambiar con el tiempo y la experiencia. Las personas usuarias pueden sentir que tienen más control a medida que se vuelven más competentes en la interacción con robots o sistemas automatizados. Esto puede influir en su actitud hacia la tecnología y su disposición para adoptarla o rechazarla. La controlabilidad también se evalúa en la herramienta UEQ+ (Schrepp & Thomaschewski, 2019).

En este contexto, la realización de estudios longitudinales en entornos de HRI industriales para evaluar cómo evolucionan las percepciones de aprendibilidad y controlabilidad a lo largo del tiempo representa una valiosa dirección de investigación. Esto podría implicar la recopilación de datos periódicos a medida que las personas trabajadoras adquieren experiencia en la interacción con robots en su entorno laboral. Los resultados de tales estudios podrían ayudar a identificar patrones y tendencias en la percepción de las personas usuarias y proporcionar conocimientos sobre cómo diseñar sistemas que se adapten mejor a las necesidades cambiantes.

Además, se podría investigar cómo los diseñadores pueden influir en la aprendibilidad y controlabilidad mediante estrategias de diseño específicas. ¿Cómo pueden diseñarse interfaces y sistemas para maximizar la facilidad de aprendizaje y la sensación de control? Esta investigación podría conducir a guías de diseño o heurísticas efectivas y prácticas para la HRI industrial.

2.5.1.5 La Confianza como factor fundamental en la aceptación de la HRI

La confianza es un factor crítico en la HRI, especialmente cuando se trata de la integración exitosa de robots en equipos humanos en entornos industriales y de trabajo. Este aspecto plantea desafíos significativos, como la falta de experiencia o conocimientos en robótica por parte de las personas, problemas de facilidad de uso y programación, así como una generalizada desconfianza hacia los robots y sus capacidades (Argyle et al., 2021).

La confianza desempeña un papel fundamental en la aceptación de los robots y su colaboración efectiva con las personas. Para que los robots se conviertan en una parte natural de la vida cotidiana de las personas usuarias, es esencial que las personas usuarias tengan experiencias positivas y adecuadas con ellos (Boden et al., 2017; Chen et al., 2020; Kahn Jr et al., 2007; Lindblom & Alenljung, 2020). Esto implica que los robots deben cumplir con los objetivos explícitos e implícitos, actuar de manera eficiente y generar sentimientos de confianza, seguridad y comodidad en las personas usuarias (Benyon, 2019).

Coronado et al. (2022), también destacan la confianza como uno de los factores de calidad más relevantes dentro de las creencias cognitivas en la HRI. Indican que, a diferencia de los factores puramente emocionales, las creencias están recibiendo más atención en la comunidad de HRI con enfoque industrial, siendo la confianza el aspecto hedónico más común evaluado o discutido. Simões et al. (2022) analizan cómo se desarrolla y establece la confianza entre la persona y el robot, ya que también lo consideran fundamental para una colaboración exitosa. La persona necesita creer en la capacidad del robot para llevar a cabo tareas de

manera precisa, confiable y segura. En este sentido, Simões et al. (2022) examinan factores que influyen en la confianza, como la experiencia previa de la persona operaria con robots, la calidad de las interacciones pasadas y la confiabilidad demostrada por el robot. Faccio et al. (2022) consideran una serie de factores humanos clave en la HRI, entre los que se encuentra la confianza, de la situación y transparencia, velocidad de movimiento y recogida del robot).

Se observa, por otro lado, que la confianza tan solo se ha medido en dos estudios experimentales identificados, concretamente en el estudio llevado a cabo por Daniel et al. (2013) y el de Akallin et al. (2022), en ambos lo estudian mediante cuestionarios. Sin embargo, esta falta de medición de la confianza en estudios experimentales es un dato interesante. Esto podría ser un área de oportunidad para futuras investigaciones. Incluir la identificación de factores que influyen en la confianza, cómo evoluciona con la experiencia y la exposición al robot, y cómo se relaciona con otros aspectos de la aceptación y la interacción. Además, la creación y validación de herramientas específicas para medir la confianza en entornos industriales podría ser un paso importante para avanzar en esta área de investigación.

2.5.1.6 La satisfacción de las personas usuarias como componente clave de la usabilidad y la hedonomía en la HRI

La definición de usabilidad según la norma ISO 9241-11 (2018) enfatiza tres componentes clave: eficacia, eficiencia y satisfacción. Estos elementos son fundamentales en el HCD y en la evaluación de la UX. Sin embargo, es interesante destacar la distinción entre factores objetivos orientados al rendimiento y aspectos subjetivos, en particular, la satisfacción.

La satisfacción se refiere a un aspecto subjetivo y emocional de la usabilidad. Se relaciona con cómo se sienten las personas usuarias al interactuar con un sistema, producto o servicio, en este caso robot. A menudo, la satisfacción está influenciada por factores emocionales, como la estética, el placer y la comodidad, que pueden ser difíciles de medir de manera objetiva.

Esta distinción entre aspectos objetivos y subjetivos es esencial porque resalta la necesidad de abordar tanto los resultados medibles como las experiencias emocionales en la evaluación. Si bien la eficacia y la eficiencia son cruciales para el rendimiento de un sistema, la satisfacción del usuario también desempeña un papel fundamental en la aceptación y la adopción a largo plazo del robot. La satisfacción puede influir en la percepción general de una persona sobre la utilidad y la facilidad de uso de un robot, lo que a su vez puede afectar su actitud y su intención de uso.

En este contexto, la hedonomía se enfoca en aspectos más positivos de las interacciones laborales al "promover la aparición de interacciones satisfactorias, que se pueden comprobar u observar" (Oron-Gilad & Hancock, 2017). Este paradigma centrado en la satisfacción y en lo afectivo, propuesto en las disciplinas de la hedonomía, contrasta con el enfoque predominantemente orientado a la seguridad y la productividad de la investigación tradicional en ergonomía.

La satisfacción se ha evaluado en dos estudios, concretamente, en los de Lasota & Shah (2015) y Almeida et al. (2020). En el primero, los autores emplearon un cuestionario que evaluaba la satisfacción con los robots como compañeros de equipo; en el segundo, se utilizó un cuestionario basado en el Cuestionario de Satisfacción de Usabilidad de Computadoras IBM.

Por lo tanto, la introducción de la hedonomía, con un enfoque centrado en la satisfacción y las experiencias emocionales en la HRI, aporta un enfoque holístico a la investigación de la percepción de las personas. Este enfoque contrasta significativamente con la tradicional orientación hacia la seguridad y la eficiencia en la investigación en robótica. El

bienestar emocional desempeña un papel influyente en la adopción de sistemas robóticos. Hasta ahora, la evaluación de la satisfacción se ha basado principalmente en cuestionarios, pero existe una oportunidad de investigar enfoques para medir la satisfacción de manera objetiva, posiblemente mediante el uso de dispositivos fisiológicos.

2.5.1.7 El ajuste entre las características que componen la interacción como base fundamental para la adopción del robot

En un contexto cada vez más orientado hacia la interacción entre personas y robots en entornos industriales, el ajuste adecuado entre las características que definen la interacción emerge como un pilar fundamental para lograr la adopción exitosa de esta tecnología. La importancia de esta adecuación se ve respaldada por un conjunto de estudios que destacan la necesidad de considerar minuciosamente una serie de factores clave al diseñar las interacciones entre personas y robots, con el propósito de optimizar la aceptación y la eficacia de estos sistemas en contextos industriales.

En este contexto, el HRCAM (Bröhl et al., 2019) enfatiza la necesidad de prestar especial atención a las variables de ajuste en el diseño de la HRI, considerando aspectos ergonómicos y la experiencia del usuario. Por otro lado, Meissner et al. (2020) reconocen la existencia de factores secundarios relacionados con el objeto, el sujeto y el contexto, que también contribuyen a la actitud de las personas hacia la aceptación de los sistemas robóticos. Si bien este enfoque proporciona una visión general de los factores influyentes, es esencial analizar críticamente la interacción compleja entre estos factores y cómo impactan en la experiencia del usuario en entornos industriales específicos.

Adicionalmente, el trabajo de Von Saucken et al. (2014) subraya la distinción entre el microcontexto y el macrocontexto en los entornos laborales, enfatizando que las condiciones específicas del entorno laboral pueden ejercer una influencia sustancial en la interacción entre las personas y robots.

El estudio de Faccio et al. (2022), por su parte, ofrece una visión global de la robótica colaborativa, considerando tres dimensiones principales: las características de los robots, las características de los sistemas de producción modernos y los factores humanos. La interacción directa entre personas y robots puede ser tanto una ventaja como una limitación en los sistemas colaborativos, dependiendo de cómo afecte a factores humanos como la ergonomía y el estrés mental.

Una oportunidad de investigación radica en desarrollar un conjunto de heurísticas específicas que garanticen o permitan evaluar el ajuste adecuado entre las características en la interacción entre personas y robots en entornos industriales. Estas heurísticas podrían proporcionar una guía detallada y práctica para mejorar la adopción y el rendimiento de sistemas robóticos en dichos contextos.

2.5.2 Métodos y medidas de evaluación de los factores humanos que influyen en la aceptación de la HRI en entornos industriales

La evaluación de los factores humanos que impactan en la aceptación de la HRI en entornos industriales involucra una amplia gama de métodos y medidas, que abarcan desde la dimensión subjetiva hasta la dimensión objetiva de la interacción. La elección y combinación adecuadas de estos enfoques dependen de los objetivos específicos de la investigación y la profundidad deseada en el análisis de los factores humanos en este contexto.

En este sentido, la evaluación de los factores humanos que influyen en la aceptación de la HRI en entornos industriales es un componente esencial para garantizar el éxito y la eficacia de esta colaboración. Sin embargo, en la evaluación de la HRI en contextos industriales, se identifica una carencia: la ausencia de un método estructurado para llevar a cabo evaluaciones efectivas. Prati et al. (2021) proponen un enfoque centrado en la UX para investigar el diálogo persona-robot, con un proceso iterativo que coloca a la persona usuaria en el centro del diseño. Sin embargo, este método presenta limitaciones al no proporcionar directrices claras para llevar a cabo evaluaciones estructuradas.

Coronado et al. (2022) ofrecen una visión panorámica de los factores, métricas y medidas empleados en la comunidad de la robótica para evaluar tanto el rendimiento como la calidad del bienestar de las personas en contextos de HRI. En su análisis, destacan un punto relevante: la mayoría de las herramientas destinadas a evaluar los factores relacionados con personas requieren técnicas de observación. Esto subraya la necesidad de incorporar la monitorización en estos procesos para lograr evaluaciones más precisas, minimizando así cualquier sesgo potencial derivado de la subjetividad del evaluador y los errores humanos. En conclusión, existe una clara oportunidad de investigación en el desarrollo de métodos estructurados y específicos que aborden los aspectos mencionados en el apartado 2.5.1 Factores humanos influyentes en la aceptación de los robots por parte de las personas en los contextos industriales

2.5.2.1 *Análisis comparativo y limitaciones identificadas en los estudios experimentales*

Dado que no se han identificado métodos estructurados de evaluación en la Interacción Humano-Robot (HRI) que aborden los aspectos previamente mencionados, se ha llevado a cabo un análisis comparativo de los estudios experimentales con el objetivo de identificar similitudes y diferencias.

La SLR ha permitido identificar 15 estudios experimentales en el ámbito de la HRI. En términos generales, los experimentos constan de tres fases: (i) antes de la ejecución de la tarea, (ii) durante la ejecución de la tarea y (iii) después de la ejecución de la tarea. El rendimiento es el factor más evaluado, principalmente a través del tiempo de ejecución de la tarea. En cuanto a la evaluación subjetiva, los cuestionarios fueron las herramientas más utilizadas, aunque se identificaron diferentes cuestionarios en los diversos estudios de caso. Sería interesante complementar el estudio con herramientas tradicionales, como cuestionarios o entrevistas, que proporcionen una visión subjetiva de las percepciones de las personas usuarias.

En la Tabla 27 se muestra un estudio comparativo de los estudios que adoptan un enfoque experimental para evaluar factores humanos en la HRI. Además de los estudios previamente identificados en la SLR, se incluyen investigaciones adicionales que reflejan una creciente tendencia hacia la incorporación de medidas fisiológicas en este campo. Se abordan un total de 17 estudios realizados en un período que abarca desde 2013 hasta 2023. Estos estudios representan una amplia gama de investigaciones que se han llevado a cabo a lo largo de la última década, lo que nos permite observar tanto la evolución como las tendencias en el campo de la HRI durante este período de tiempo. La heterogeneidad identificada entre los

diversos estudios sugiere la necesidad de una mayor estandarización y la exploración de enfoques más integrales para evaluar la HRI en entornos industriales.

De los 17 estudios analizados, 16 de ellos incorporan al menos un cuestionario como parte de sus métodos de evaluación. El único estudio que no utiliza un cuestionario en su enfoque opta por realizar una entrevista, dado que cuenta con un único usuario. Dentro de la variedad de cuestionarios empleados, destaca el NASA-TLX, utilizado en 4 de los estudios. Esta elección frecuente del NASA-TLX sugiere que es una herramienta reconocida y ampliamente validada para medir la carga de trabajo percibida y otros aspectos relacionados con la HRI. No obstante, es importante señalar que el NASA-TLX no aborda aspectos hedonómicos de la experiencia de la interacción con un robot, lo que resalta la necesidad de identificar o desarrollar un cuestionario más completo que permita una evaluación integral de la percepción de las personas usuarias.

En cuanto a las mediciones fisiológicas, se observa que 4 de los estudios han incorporado este tipo de evaluaciones en sus metodologías. Destaca la particularidad de que 3 de estos estudios son notoriamente actuales, con fechas de publicación en los años 2022 y 2023, lo que subraya el hecho de que los avances tecnológicos y el desarrollo de wearables más sofisticados están posibilitando la incorporación de la monitorización fisiológica en la investigación. Sin embargo, es importante destacar que las medidas fisiológicas registradas se centran principalmente en el ECG y la EDA. Hasta el momento, no se ha identificado ningún estudio que haya utilizado la EEG en conjunto con estas mediciones fisiológicas.

En lo que respecta a las medidas de desempeño, se ha observado que 8 de los 17 estudios incorporan evaluaciones relacionadas con el rendimiento. La métrica de desempeño más comúnmente utilizada en estos estudios es el "tiempo de ejecución de la tarea". Esta medida se utiliza para evaluar la eficiencia con la que se realizan las tareas durante la interacción entre personas y robots.

Sin embargo, es importante destacar que la elección de las métricas de desempeño puede variar según la tarea específica que se esté evaluando en cada estudio. Esto implica que los indicadores de desempeño pueden ser altamente contextuales y depender en gran medida de la naturaleza de la tarea en cuestión. Dado que la tarea desempeña un papel fundamental en la determinación de la calidad de la interacción HRI, es esencial contar con indicadores de desempeño que reflejen con precisión la eficiencia y efectividad de dichas interacciones.

En este sentido, la diversidad de indicadores de desempeño observados en estos estudios es un reflejo de la necesidad de adaptar las medidas a las características específicas de las tareas realizadas en diferentes contextos de HRI. Esta flexibilidad en la selección de indicadores permite una evaluación más completa y precisa de la calidad de la interacción entre personas y robots en una variedad de situaciones y escenarios.

Por todo ello, las futuras experimentaciones deben integrar un enfoque holístico para capturar las percepciones de las personas en todo momento, es decir, antes, durante y después de una interacción mediante enfoques próximos a Roto et al. (2011). Incluir diferentes métodos de medición en diferentes momentos de la interacción, tanto cuantitativos como cualitativos, objetivos y subjetivos, e integrar aspectos de género para garantizar la interpretación de los datos y la comprensión del flujo completo de la interacción. Ante la carencia de experimentaciones que evalúen el carácter temporal de la UX, es decir, el antes, el durante y el después de la interacción, existe la oportunidad de investigar cómo considerarla en las tres fases.

Tabla 27: Análisis comparativo de los estudios que adoptan un enfoque experimental para analizar los factores humanos en la HRI industrial.

Año	Autor/es	Mediciones y métricas relacionadas con factores humanos		Mediciones y métricas relacionadas con el desempeño	Contribución principal
		Evaluación de la percepción	Evaluación fisiológica		
2013	Schillaci et al.	<ul style="list-style-type: none"> Cuestionario Godspeed 	-	-	Identificación de las características físicas y de comportamiento que deben destacarse para garantizar una buena UX en la HRI.
2013	Daniel et al.	<ul style="list-style-type: none"> Preguntas abiertas sobre la confianza en la automatización. Entrevista con preguntas adaptadas de WAI Site Usability Testing Questions sobre las impresiones sobre la UI actual y el funcionamiento de la célula robótica. Cuestionario de Condiciones de Flujo 	-	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Número de interacciones realizadas Relación entre las interacciones con la pantalla táctil y el teclado 	Mejorar la HRI mediante la implementación de una UI.
2013	Harriott et al.	<ul style="list-style-type: none"> NASA-TLX 	<ul style="list-style-type: none"> Bioharness ECG (HR, HRV, RPM, ST) 	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea 	Evaluación de la carga de trabajo en equipos humano-humano y humano-robot que realizan las mismas tareas.
2015	Lasota & Shah	<ul style="list-style-type: none"> Cuestionario (Satisfacción con el robot como compañero; y seguridad y comodidad percibida) 	-	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Movimiento simultáneo Distancia media de separación Tiempo de inactividad del robot Tiempo de inactividad de la persona 	Análisis de la respuesta humana a distintos niveles de adaptación del robot en tareas colaborativas.
2017	Danielsson et al.	<ul style="list-style-type: none"> Cuestionario SUS Cuestionario de interés general 	-	-	Evaluación de la eficacia de las instrucciones en RA para el ensamblaje colaborativo entre humanos y robots.
2018	Aromaa et al.	<ul style="list-style-type: none"> NASA-TLX 	-	-	Identificación de las características físicas y de comportamiento esenciales para asegurar una UX óptima en la HRI a través del uso de DHM en la evaluación de prototipos de interfaces gestuales.
2018	Tang & Webb	<ul style="list-style-type: none"> Cuestionario SUS 	-	-	Evaluación para abordar problemas ergonómicos en el diseño y comparación de la UX de este sistema con un controlador de aprendizaje convencional en el estudio.
2019	Kidal et al.	<ul style="list-style-type: none"> Entrevista 	-	-	Evaluación piloto cualitativa en la que un trabajador de montaje con discapacidades cognitivas y expertos multidisciplinares interactuaron con un robot colaborativo en

Año	Autor/es	Mediciones y métricas relacionadas con factores humanos		Mediciones y métricas relacionadas con el desempeño	Contribución principal
		Evaluación de la percepción	Evaluación fisiológica		
2020	Almeida et al.	<ul style="list-style-type: none"> IBM Computer Usability Satisfaction Questionnaire Cuestionario basado en las preguntas de presencia de Slater et al. (1994) y Usoh et al. (2000). 	-	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Errores Precisión de seguimiento de la trayectoria 	<p>tareas de montaje rutinarias.</p> <p>Análisis de la transferencia habilidades y experiencia a lugares remotos a través de sistemas de teleoperación.</p>
2020	Hietanen et al.	<ul style="list-style-type: none"> Cuestionario autogenerado para evaluar la UX compuesto por los subconstructos seguridad, procesamiento de la información, ergonomía, autonomía, competencia y afinidad. 	-	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Tiempo de inactividad del robot 	<p>Modelo basado en sensores de profundidad para la supervisión del espacio de trabajo y una UI interactiva de RA para una HRC segura en la fabricación industrial.</p>
2020	Pantano et al.	<ul style="list-style-type: none"> NASA-TLX 	-	-	<p>Evaluación de la carga de trabajo percibida por las personas usuarias al relacionarse con el robot y su interfaz.</p>
2020	Beschi et al.	<ul style="list-style-type: none"> Cuestionario autogenerado sobre la peligrosidad percibida 	-	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Distancia media de separación Cociente entre el tiempo necesario para completar la tarea con y sin el robot 	<p>Análisis de los parámetros de planificación de movimiento en la HRC.</p>
2021	Colim et al.	<ul style="list-style-type: none"> Cuestionario para evaluar las percepciones de las personas trabajadoras sobre los robots en la industria, y las mejoras ergonómicas asociadas. 	-	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Variabilidad en los tiempos de producción Índice de producción 	<p>Evaluación del riesgo musculoesquelético asociado al puesto de premontaje antes y después de la implementación del control HRC, así como el análisis de los indicadores de bienestar y aceptación de la robótica, basándose en las percepciones de las personas trabajadoras.</p>
2021	Joosse et al.	<ul style="list-style-type: none"> Cuestionario Godspeed 	-	-	<p>Análisis de las respuestas de las personas a las violaciones del espacio personal cometidas por los robots en comparación con las de las personas y determinar los efectos de las características de los robots (sonido, tamaño y velocidad) en la aceptación.</p>
2022	Gervasi et al.	<ul style="list-style-type: none"> Cuestionario NARS Cuestionario sobre la calidad de la interacción SAM 	<ul style="list-style-type: none"> Skin Conductance Responce 	<ul style="list-style-type: none"> Tiempo de ejecución de la tarea Velocidad del robot Distancia entre la persona y el robot 	<p>Evaluación experimental sobre las relaciones entre la velocidad, distancia y control de la HRC, la calidad de interacción</p>

Año	Autor/es	Mediciones y métricas relacionadas con factores humanos		Mediciones y métricas relacionadas con el desempeño	Contribución principal
		Evaluación de la percepción	Evaluación fisiológica		
			(SCR)		percibida, el estado afectivo y el estrés.
			<ul style="list-style-type: none"> • HRV • EDA • HRV • IBI 	-	Relación entre factores como la comodidad, la previsibilidad, la sensación de control y la confianza, y la seguridad percibida en la interacción entre humanos y robots utilizando medidas subjetivas y objetivas.
2022	Akalin et al.	<ul style="list-style-type: none"> • Cuestionario sobre el confort basado en el estudio de (Kim & Mutlu, 2014). • Cuestionario sobre la seguridad percibida (Akalin et al., 2019). • Cuestionario sobre el sentido de control adaptado de Strube & Werner (1984). • Cuestionario Trust Perception Scale-HRI (Schaefer, 2016). • Cuestionario post ejecución de Weiss et al. (2009). • SAM 	<ul style="list-style-type: none"> • HRV • EDA • HRV • IBI 	-	Relación entre factores como la comodidad, la previsibilidad, la sensación de control y la confianza, y la seguridad percibida en la interacción entre humanos y robots utilizando medidas subjetivas y objetivas.
2023	Gervasi et al.	<ul style="list-style-type: none"> • NASA-TLX • SAM 	<ul style="list-style-type: none"> • HRV • EDA 	-	Evaluación experimental que emula los turnos de trabajo e investiga las diferencias entre los procesos de ensamblaje repetitivo manual y los procesos de ensamblaje repetitivo.

Los estudios identificados en la Tabla 27 utilizan principalmente cuestionarios como herramienta de evaluación de UX en HRI, que, a pesar de su amplia aplicación, es importante destacar que la evaluación retrospectiva y subjetiva proporcionada por los cuestionarios y entrevistas no siempre constituye el enfoque óptimo. Uno de los principales desafíos de este enfoque radica en que se basa en la memoria y la percepción subjetiva de la persona usuaria, lo que puede dar lugar a errores debido a recuerdos inexactos o sesgos cognitivos. Para abordar esta limitación, resulta especialmente interesante considerar la incorporación de la monitorización fisiológica como una herramienta complementaria. De acuerdo con Neumann & Westbury (2011), las medidas fisiológicas se consideran más objetivas en comparación con las medidas basadas en autorregistro. Esta objetividad se debe al hecho de que las respuestas fisiológicas no dependen de la memoria o la percepción subjetiva de la persona usuaria, sino que son una manifestación directa de sus reacciones fisiológicas en tiempo real. Además, es importante destacar que las oportunidades para utilizar señales fisiológicas en la evaluación de la UX están en constante crecimiento gracias a los avances en la tecnología de sensores y el procesamiento de señales (Oliveira & Guimarães, 2013). Estos avances han facilitado la incorporación de medidas fisiológicas más precisas y menos invasivas en los estudios de HRI. Por lo tanto, la monitorización fisiológica ofrece la ventaja de proporcionar una visión más objetiva y detallada de la experiencia de la persona usuaria durante la interacción con robots, lo que puede enriquecer significativamente la comprensión de la UX en este contexto.

Por último, destaca que, en los estudios experimentales analizados, ha sido evidente la falta de atención a la variable de género como parte integral de la investigación. Esta omisión representa una oportunidad significativa para enriquecer la presente investigación, ya que incorporar la perspectiva de género se considera fundamental para abordar y comprender la brecha digital de género existente en el campo de la HRI. La inclusión de la perspectiva de género en esta investigación puede proporcionar información valiosa sobre cómo las diferencias de género pueden influir en la experiencia de la persona usuaria en las interacciones con robots. Esto es particularmente relevante en un contexto en el que la tecnología y la robótica desempeñan un papel cada vez más importante en la sociedad, y es esencial garantizar que los beneficios y desafíos asociados se distribuyan equitativamente entre diferentes grupos de población, independientemente de su género. Además, la consideración de la perspectiva de género en esta investigación se alinea con las políticas europeas propuestas en el marco de *Horizon Europe* con respecto a la igualdad de género. Estas políticas promueven la equidad de género como un objetivo central en la investigación y la innovación, reconociendo que es esencial abordar las disparidades de género en todas las áreas de la tecnología, incluida la interacción con robots.

2.5.2.2 Análisis comparativo de los métodos y métricas de evaluación en HRI en entornos industriales

En líneas generales, la literatura tiende a estar predominantemente orientada hacia cuestiones técnicas y soluciones tecnológicas (Prati et al., 2021). El enfoque "centrado en el robot" puede descuidar factores humanos esenciales para la implementación exitosa de sistemas robóticos en entornos reales. Para abordar esta brecha, es crucial adoptar un enfoque centrado en la persona, comprendiendo sus necesidades, expectativas y preferencias para un diseño compatible con las capacidades humanas y mejorar la productividad. El análisis de las variables de la HRI es de gran importancia en la robótica, ya que permite comprender de manera integral la dinámica de las interacciones entre personas y robots. Al examinar factores como la comunicación, la asignación de tareas y la fiabilidad entre personas y robots, los investigadores pueden identificar posibles desafíos y oportunidades para mejorar la colaboración y aumentar el rendimiento. Además, dicho análisis puede informar el desarrollo de estrategias efectivas de colaboración y permitir el diseño de sistemas robóticos que se ajusten mejor a las necesidades y preferencias humanas.

En un esfuerzo por cuantificar la HRI, Kokotinis et al. (2023) han desarrollado un índice medible que se calcula a partir de indicadores individuales destinados a reflejar varios aspectos, incluyendo el rendimiento, costo, calidad y flexibilidad de la aplicación de la HRI. Este índice implica la cuantificación de métricas que abarcan tres categorías distintas: métricas relacionadas con el robot, métricas relacionadas con la persona y métricas relacionadas con la HRI en sí. La implementación de esta herramienta resulta sencilla y eficaz, proporcionando un indicador conciso que permite una comprensión general de la evaluación de la HRI, aunque cabe destacar que se trata de una evaluación inicial que proporciona una visión general.

En la misma línea Grandi et al. (2022) proponen la creación de UXI para evaluar la calidad de la interacción humano-sistema durante las tareas laborales y, en consecuencia, evaluar tanto el proceso como el espacio de trabajo. Integra cuatro clases de datos para objetivar la UX y la carga de trabajo percibida: análisis postural basado en el método REBA, datos fisiológicos basados en la HR y PD, rendimiento basado en el tiempo de ejecución de tarea y datos subjetivos basados en el cuestionario NASA-TLX. A partir de estos datos, los autores proponen una serie de fórmulas matemáticas que se combinan para calcular el indicador UXI. En consonancia con el enfoque propuesto por Grandi et al. (2022), Khamaisi, Grandi, et al. (2022) se embarcan en un esfuerzo para perfeccionar el UXI. El índice UXI se compone de la consideración de varios parámetros clave para su cálculo. Estos parámetros incluyen la Carga Mental de Trabajo (MWL), que se deriva de la interacción compleja entre los requisitos de una tarea dada, las circunstancias en las que se lleva a cabo y las habilidades, comportamientos y percepciones de la persona operaria. Además, se incluye el Estrés (S), que se refiere a la respuesta fisiológica del sistema nervioso simpático del individuo ante la interacción con el sistema. Finalmente, se considera la Carga Postural de Trabajo (PWL), que se evalúa a través del puntaje RULA, un índice ergonómico utilizado para evaluar la postura adoptada por la persona usuaria durante la interacción. Si bien estas aproximaciones proporcionan una evaluación integral, es importante destacar que estos datos no abordan aspectos emocionales o de naturaleza hedónica, lo que podría requerir métodos adicionales para una comprensión completa sistema desde una perspectiva emocional.

2.5.2.2.1 Medidas fisiológicas para la evaluación en la HRI

El análisis de los estudios experimentales ha puesto en evidencia el creciente interés en la incorporación de medidas fisiológicas en la evaluación de la HRI. A lo largo de los años y en consonancia con los avances tecnológicos, se ha observado un uso cada vez mayor. No obstante, en el contexto de los estudios experimentales, prevalece el uso de dispositivos de ECG y EDA. Aunque estos dispositivos proporcionan una gran cantidad de información valiosa sobre las respuestas fisiológicas de los individuos durante la HRI, es importante señalar que aún existen limitaciones en la captura de las respuestas emocionales en comparación con el EEG. El ECG mide la actividad eléctrica del corazón y puede revelar información sobre la variabilidad de la frecuencia cardíaca (HRV), que está relacionada con la regulación del sistema nervioso autónomo y puede indicar niveles de estrés o relajación. Por otro lado, el EDA mide las respuestas eléctricas de la piel y puede ser sensible a cambios emocionales, como la sudoración en respuesta al estrés. Ambos dispositivos son valiosos para comprender las reacciones fisiológicas durante la interacción humano-robot. Sin embargo, el EEG, que registra la actividad eléctrica del cerebro, puede proporcionar una visión más directa de las respuestas emocionales y cognitivas. Puede detectar patrones específicos de actividad cerebral relacionados con estados emocionales como la felicidad, el estrés o la concentración. Esto significa que el EEG puede ser especialmente útil para capturar las respuestas emocionales y cognitivas en tiempo real durante la interacción con robots, lo que ofrece una comprensión más profunda de la UX.

Por otro lado, en el marco de esta investigación realizamos una SLR (Apraiz et al., 2021) donde observamos que el EEG emerge como el dispositivo de medición más

frecuentemente utilizado en diversos campos de investigación, aunque no se ha aplicado en estudios específicos de HRI. Esto sugiere un interés genuino en la aplicación del EEG para evaluar la UX en este ámbito. Sin embargo, es importante destacar que, hasta el momento, su implementación en la HRI ha sido limitada debido a ciertas limitaciones y desafíos asociados. Una de las limitaciones clave del EEG en el contexto de HRI es el movimiento físico de la persona usuaria durante la interacción con un robot puede introducir ruido en los datos EEG. Además, en los test de UX o usabilidad, es común que la persona participante realice movimientos físicos durante la interacción. Otra limitación es la necesidad de un alto grado de conocimiento para procesar y analizar los datos EEG de manera efectiva. La interpretación de las señales cerebrales requiere experiencia en neurociencia y procesamiento de señales, lo que puede dificultar su adopción generalizada en entornos de investigación que no cuentan con especialistas en esta área. Otro desafío importante es el tiempo requerido para la calibración de los dispositivos EEG. Los electrodos deben colocarse de manera precisa en el cuero cabelludo de la persona usuaria, lo que puede ser un proceso demorado y que requiere una calibración precisa para cada individuo. Esta calibración individualizada puede ser poco práctica en situaciones de HRI, donde se busca una implementación rápida y sencilla. No se ha utilizado el EEG en ninguno de los casos de estudio realizados con robots, puesto que, además de dificultar el registro y análisis de datos, puede considerarse más molesto y limitar los movimientos de las personas participantes.

En este sentido, podría existir otra línea de investigación en relación con las emociones positivas y negativas y el uso de herramientas fisiológicas. La comprensión de las emociones humanas y su influencia en la HRI es un aspecto crítico para mejorar la calidad de esta interacción, especialmente en entornos industriales donde la eficiencia y la seguridad son fundamentales (Gualtieri et al., 2021). Esto involucra el uso de herramientas fisiológicas como sensores HR, EDA o incluso EEG para rastrear las respuestas emocionales de las personas trabajadoras durante la interacción con robots industriales.

El creciente número de implementaciones de sistemas de reconocimiento biométrico hizo imperativa la producción masiva de sistemas fisiológicos y redujo los precios (Thakkar, 2018). Esta reducción de los precios podría también estar relacionada con el aumento del uso de este tipo de tecnología en test de UX durante los últimos 5 años. Aun así, existen ciertas variables que podrían limitar este incremento. Por un lado, el precio podría ser uno de los motivos (Taffese, 2017). A pesar de la reducción del precio debido a la producción masiva, podría seguir siendo elevado para los profesionales de la UX.

El GSR, en cambio, es un dispositivo menos sensible al ruido y menos ambiguo que los sensores como EMG o HR, es por ello por lo que resulta más fácil analizar los resultados obtenidos (Bruun, 2018). En la misma línea, el *eyetracking* permite conocer el movimiento y la fijación de los ojos, pero debido a que las fijaciones no son exactamente equivalentes a la atención, se puede extrapolar, pero no siempre se puede estar seguro de los resultados (Teverovsky, s. f.). La fiabilidad de los resultados puede resultar crítica en algunos casos.

Sin embargo, el desarrollo continuo y la mejora de los dispositivos de monitorización fisiológica en términos de la complejidad y análisis de la señal, podría implicar el incremento del uso de estos dispositivos en el campo de estudio de la UX en general. Con los avances en la tecnología, los dispositivos fisiológicos podrían ser más accesibles para usar en los laboratorios de UX con frecuencia. Con un uso más frecuente debido a la disponibilidad de la tecnología, más profesionales de la UX podrían experimentar y continuar encontrando nuevas y útiles aplicaciones de la tecnología en diferentes sectores. En esta línea, el uso cada vez mayor de los *smartwatches* que incluyen sensores fisiológicos (Rawassizadeh et al., 2015) podría tener un potencial considerable como fuente de datos para medir y entender la UX durante las interacciones tanto en entornos reales como en laboratorio. Tal y como indican Maia y Furtado

(2016), el nivel de invasivo de las herramientas fisiológicas es crítico, y el ofrecido por los *smartwatches* es inferior a la mayoría de los dispositivos.

En este contexto, se abre una oportunidad en este campo mediante la combinación de diferentes herramientas fisiológicas que podría ayudar en la precisión y fiabilidad de los resultados, ya que proporcionaría un mejor entendimiento del contexto en el que se da la interacción. Además, resultaría interesante complementar el estudio con herramientas tradicionales como los cuestionarios o las entrevistas, que permiten conocer la percepción de las personas usuarias de forma subjetiva. La combinación de los diferentes métodos permitiría una evaluación más completa de la UX, es decir, una evaluación que incluya las medidas fisiológicas y las obtenidas mediante percepciones subjetivas.

- *El tamaño de la muestra en los estudios experimentales*

Trabajos como los de Virzi (1992), Lewis (1994) y Nielsen (2000) afirman que 5 participantes son suficientes para identificar el 80% de problemas de usabilidad. Sin embargo, el uso de muestras de tamaño pequeño puede producir una gran variabilidad en los resultados de las pruebas que no se pueden ajustar completamente (Cazañas et al., 2017). En el estudio realizado por Faulkner (2003) se analizaron 60 personas usuarias y se tomaron muestras de conjuntos aleatorios de 5 o más del conjunto, para demostrar los riesgos de usar sólo 5 participantes y los beneficios de usar más. Algunos de los conjuntos de 5 participantes seleccionados al azar encontraron el 99% de los problemas; otros conjuntos encontraron sólo el 55%. Con 10 personas usuarias, el porcentaje más bajo de problemas revelados por un conjunto se incrementó al 80%, y con 20 personas usuarias, al 95%.

En este estudio, se ha tratado de conocer cómo realizan los profesionales los testeos de UX con monitorización fisiológica, y así, entre otros datos, obtener una estimación del tamaño apropiado teórico. Sin embargo, los resultados muestran que no existe un tamaño de muestra ampliamente validado. Así como los test de usabilidad tienen un tamaño de muestra acotado y ampliamente discutido por la comunidad científica, los test que incluyen dispositivos fisiológicos y miden las respuestas fisiológicas no tienen una muestra validada. Se ha observado una gran variabilidad en el tamaño de las muestras de los casos de estudio identificados. Siendo así, establecer un tamaño de muestra aproximado según la herramienta fisiológica que se va a utilizar en los diferentes sectores y profundizar en los diseños experimentales para cada uno de los dispositivos, se presenta como una oportunidad de investigación clara.

2.5.2.2.2 La evaluación de la UX mediante cuestionarios en el contexto industrial

Aunque los procesos industriales están cada vez más automatizados, el desempeño humano sigue siendo el principal responsable de la calidad de los productos y la productividad de las fábricas. En este contexto, comprender cómo perciben la interacción las personas trabajadoras con los sistemas de producción y cómo experimentan el entorno de la fábrica es fundamental para modelar adecuadamente la interacción humana y optimizar los procesos (Peruzzini et al., 2017).

La evaluación de la UX en el contexto de la HRI, presenta desafíos particulares y, hasta el momento, no se han identificado cuestionarios específicos diseñados exclusivamente para esta área. Esto refleja una brecha en la literatura y las herramientas disponibles para medir y comprender la calidad de la UX en interacciones con robots. Dado su carácter interdisciplinario, la HRI a menudo incorpora y adapta herramientas y enfoques provenientes de disciplinas afines, como HCI o la ergonomía. En el análisis de los estudios experimentales se observa la inclusión de cuestionarios como NASA-TLX que, si bien no está específicamente diseñado para evaluar la interacción con robots, se centra en aspectos relacionados con la carga de trabajo percibida por las personas usuarias. No obstante, es importante señalar que

el NASA-TLX podría no captar completamente la UX en su totalidad. Este cuestionario tiende a centrarse más en aspectos de carga de trabajo y eficiencia, dejando de lado elementos más hedonómicos relacionados con la satisfacción y la experiencia emocional de la persona usuaria al interactuar con el robot. Esto significa que, aunque el NASA-TLX puede proporcionar información valiosa sobre el aspecto de carga mental de trabajo, podría no ser suficiente para evaluar la percepción global y emocional de la persona usuaria con respecto al robot. En los estudios experimentales también se ha observado la aplicación del cuestionario Godspeed (Bartneck et al., 2008) en dos ocasiones y mide el antropomorfismo, la animación, la simpatía, la inteligencia y la seguridad percibidas de los robots. En una ocasión también se ha utilizado el cuestionario SUS (Brooke, 1996), que mide la usabilidad y la aprendibilidad. Además, los estudios experimentales incluyen cuestionarios no estandarizados pero que miden constructos como la confianza, la satisfacción, la seguridad y comodidad percibida, estrés físico y mental y el riesgo percibido.

Sin embargo, cuestionarios más completos y estandarizados en la evaluación de la UX como puede ser el cuestionario UEQ, aun no se han aplicado en HRI. Esta carencia subraya la necesidad de desarrollar y aplicar herramientas de evaluación específicas para comprender de manera más precisa y completa la calidad de las interacciones entre personas y robots. En lugar de crear cuestionarios completamente nuevos, se podría considerar la adaptación y extensión de cuestionarios existentes, como el UEQ, para abordar dimensiones específicas de la HRI.

En la revisión de la literatura presentada en el apartado 2.2 Modelos de evaluación en HRI: Revisión Sistemática de la Literatura (SLR), se ha identificado, además, el cuestionario SE-HRI (von der Pütten & Bock, 2018), un cuestionario que evalúa la percepción de autoeficacia de las personas usuarias en la relación persona-robot. En este sentido, se observa la aplicabilidad de este cuestionario en el contexto de la HRC industrial, ya que podría tener encaje.

Además, los factores humanos evaluados identificados en los estudios experimentales identificados también pueden ser evaluados con los cuestionarios de la UX estandarizados. En este sentido, la confianza, podría medirse con SUPR-Q (Sauro, 2015) o UEQ+ (Schrepp & Thomaschewski, 2019), la satisfacción podría medirse, por ejemplo, con el cuestionario USE (Lund, 2001) o UMUX (Finstad, 2010). La aprendibilidad, además de con el cuestionario SUS (Brooke, 1996), también podría medirse con QUIS (Chin et al., 1988), SUMI (Kirakowski & Corbett, 1993), PUTQ (Lin et al., 1997), WAMMI (Kirakowski & Cierlik, 1998), USE (Lund, 2001), UEQ (Laugwitz et al., 2008), AltUsability (J. R. Lewis et al., 2015) y UEQ+ (Schrepp & Thomaschewski, 2019). Así como la usabilidad, que se podría evaluar con los cuestionarios Perceived Usefulness and Ease of Use (Davis, 1989), CSUQ (J. R. Lewis, 1992a), PSSUQ (J. R. Lewis, 1992b), USE (Lund, 2001), AltUsability (Schrepp & Thomaschewski, 2019) y UEQ+ (Schrepp & Thomaschewski, 2019).

Ante este escenario se observa que los cuestionarios son herramientas capaces de evaluar la UX, fáciles de aplicar y que no suponen grandes recursos. A pesar de no haber identificado ningún cuestionario diseñado específicamente para entornos industriales de HRI, los cuestionarios de UX se presentan como una gran oportunidad, debido a que recogen constructos que pueden ser interesantes para aplicar en el contexto de la nueva industria y conocer así la percepción de las personas usuarias.

2.5.3 La influencia de las variables cinemáticas en la percepción de las personas usuarias

Las variables cinemáticas del robot, como la velocidad y la aceleración, desempeñan un papel esencial en la forma en que las personas perciben un robot (Arai et al., 2010; Dehais

et al., 2011; Dragan et al., 2015; Hanajima et al., 2017; Lasota & Shah, 2015). La influencia de las variables cinemáticas en la percepción de las personas usuarias en el contexto de la HRI es un área de investigación que ha cobrado una creciente relevancia en los últimos años (Joosse et al., 2021). La manera en que los robots se mueven, actúan y se comportan en presencia de las personas usuarias desempeña un papel fundamental en la calidad de la interacción y en la percepción que estas tienen de la tecnología robótica. Sin embargo, la velocidad y aceleración de los movimientos se han considerado principalmente en robots colaborativos por razones de seguridad. Esto hace que los investigadores y desarrolladores pasen por alto las respuestas emocionales de la persona operaria al especificar las variables cinemáticas del robot (Koppenborg et al., 2017).

Koppenborg et al. (2017) estudiaron los efectos de la velocidad de movimiento y la previsibilidad de la trayectoria en las personas operarias humanas utilizando un entorno de realidad virtual. Descubrieron que los movimientos más rápidos resultaban en calificaciones más altas de carga de trabajo y ansiedad. En los estudios de Lasota & Shah (2015) y Beschi et al. (2020) también exploraron los efectos del movimiento. Sin embargo, no exploraron de manera integral los aspectos cognitivos y perceptuales. Rubagotti et al. (2021) identificaron los principales factores relacionados con el movimiento y las características del robot que determinan la percepción de seguridad. Entre otras cosas, concluyeron que cuanto menor es la velocidad de los robots, más seguros se perciben.

Según Joosse et al. (2021), una posible razón de la preferencia de la gente por una velocidad más lenta de los robots puede ser que no todos los robots dan una indicación clara de cuándo van a detenerse, es decir, no proporcionan *feedback* de su intencionalidad. Una forma de hacerlo es equipar a los robots con *feedbacks* funcionales; para transmitir las intenciones. En este sentido, hay una oportunidad de investigación con respecto a las variables cinemáticas y la correlación en la percepción de las personas usuarias y su desempeño.

En este sentido, se abre una oportunidad de investigación centrada en explorar el impacto de las variables cinemáticas de un robot en la percepción de las personas usuarias en entornos de HRI, teniendo en cuenta también aspectos de naturaleza hedonómica. Por otro lado, se podría tratar de comprender cómo la capacidad de un robot para comunicar claramente sus intenciones a través de su comportamiento cinemático influye en la percepción de las personas usuarias. Finalmente, también sería interesante identificar los aspectos cognitivos y perceptuales más relevantes relacionados con las variables cinemáticas de los robots y su efecto en la percepción de seguridad.

2.5.4 Optimización de la UX en la HRI a través del *feedback* de las interfaces

En la investigación sobre la HRI, la percepción de la UX es un concepto subjetivo que varía entre personas. Lindblom & Alenljung (2020) destacan que, en este contexto, para mejorar la percepción de la UX, es fundamental que las personas sean capaces de comprender al robot, anticipar sus intenciones, prever secuencias de acciones, evaluar decisiones, determinar la necesidad de intervención, definir una secuencia de acciones y ejecutarla. En este proceso, la interfaz actúa como el canal de comunicación central entre las partes involucradas.

Uno de los aspectos cruciales que más afecta a la UX en HRI es el *feedback* proporcionado a las personas a través de la interfaz (Cohen et al., 2018). Este *feedback* es esencial para que la interacción sea percibida de manera satisfactoria por las personas usuarias (Schillaci et al., 2013; Kildal et al., 2019). Cohen et al. (2018) resaltan la importancia de seleccionar el modo de reacción adecuado, incluyendo el canal, la frecuencia y la intensidad del *feedback*, basándose en una observación y análisis exhaustivos de los elementos involucrados.

Las personas usuarias han de ser conscientes de la situación y entender el comportamiento del sistema, para así poder intervenir en situaciones dinámicas e imprevistas. Esto implica, tal y como indican Villani et al. (2018), diseñar interfaces intuitivas y un *feedback* adecuado como parte integral de garantizar la seguridad en la HRI. En este sentido, la robótica afectiva podría ser adecuada a fin de garantizar una interfaz intuitiva, aliviando la carga cognitiva de la persona usuaria ya que el robot se adaptaría a la situación de la persona (Villani et al., 2018).

Por lo tanto, se concluye que, para establecer una comunicación bidireccional efectiva entre persona y robot, es esencial proporcionar un *feedback* adecuado. No obstante, este contexto también plantea preguntas de investigación, como la naturaleza del *feedback* en términos de acciones del robot y emociones del usuario. Se abre la puerta a preguntas como: ¿Qué tipo de interfaz es más adecuada en cada caso (visual, voz, háptica, etc.)? ¿Se percibiría mejor una interfaz unimodal o una multimodal? En caso de ser multimodal, ¿cuál sería la combinación óptima? ¿Cómo afecta este *feedback* a las emociones de las personas usuarias?

2.5.4.1 *La multimodalidad y la combinación óptima de las diferentes modalidades en las interfaces en la HRI*

La HRI ha emergido como un campo de estudio esencial en la era contemporánea, en la que la colaboración entre personas y robots se ha vuelto cada vez más prominente en diversos contextos industriales y sociales. La interacción efectiva entre estas dos entidades, personas y robots, es fundamental para lograr una colaboración armoniosa y productiva. En este contexto, las interfaces juegan un papel central como canales primarios de comunicación, sirviendo como mediadores esenciales entre las personas usuarias y los sistemas robóticos.

Las personas operarias deben poder proporcionar instrucciones simples a los robots sin distracciones significativas de sus tareas principales. De manera recíproca, los robots deben suministrar información clara a las personas usuarias, permitiendo una comprensión inmediata e interpretación de los datos intercambiados. Este requisito adquiere aún mayor relevancia en el contexto de colaboraciones más cercanas entre personas y robots, donde la facilidad y naturalidad de la interacción son esenciales. Niveles crecientes de interacción entre personas y robots requieren interfaces más avanzadas para garantizar la efectividad y la naturalidad (Prati et al., 2020).

En esta línea, el establecimiento de una comunicación efectiva entre personas y robots a través de interfaces adecuadas es crucial para el éxito de la colaboración. Para lograrlo, se deben definir tanto las interacciones deseadas entre personas y robots como el propósito de los intercambios de información. Estos elementos están en gran parte delineados por el alcance de la aplicación y las funciones específicas de las personas y los robots, y deben ser adaptados a diferentes contextos para asegurar la eficacia.

La multimodalidad de estas interfaces añade un desafío fundamental a la búsqueda de una colaboración efectiva ya que agrega un nivel de complejidad a la comunicación. Las interfaces multimodales deben ser coherentes y complementarias entre sí para que la comunicación sea efectiva. Evaluar cómo diferentes modalidades se integran y se apoyan mutuamente en la HRI es esencial para comprender su impacto en la experiencia de la persona usuaria. La efectividad de las interfaces multimodales en la HRI se mide en parte por su capacidad para permitir interacciones más naturales entre personas y robots. Evaluar si las interfaces logran este objetivo es un desafío importante.

Asimismo, las interfaces persona-robot pueden generar diferentes tipos de interacciones. La comunicación bidireccional efectiva entre personas y robots en entornos industriales requiere una atención especial al *feedback* proporcionado. Cohen et al. (2018), Schillaci et al. (2013) y Kildal et al. (2019) subrayan la importancia de brindar un *feedback*

adecuado en la HRI. Según la tipología de la comunicación, las interfaces persona-robot pueden clasificarse en cuatro categorías principales (Prati et al., 2021): (i) pantallas visuales (por ejemplo, GUI, interfaces de RA), (ii) gestos (por ejemplo, movimientos faciales y de la mano), (iii) habla y lenguaje natural (por ejemplo, respuestas auditivas y basadas en texto), e (iv) interacciones físicas y hápticas. El sector industrial actualmente se inclina más por las interfaces visuales (Prati et al., 2021). Esto se debe, probablemente, a la multiplicidad de opciones de información que se ofrecen al robot y, en consecuencia, a la variedad de retroalimentación que se recibe del mismo. Conforme aumenta el nivel de interacción, también lo hace el riesgo asociado, especialmente en escenarios de colaboración estrecha entre personas y robots, donde la seguridad es de máxima importancia.

El desarrollo de soluciones efectivas en el ámbito de la HRI implica la necesidad de un marco analítico y de síntesis que incluya la clasificación y caracterización del problema, junto con pautas y directrices para desarrollar soluciones que se integren de manera coherente con los requisitos de producción existentes.

La UX también desempeña un papel central en la calidad de los sistemas interactivos. Más allá de la usabilidad, la UX abarca las emociones antes, durante y después de la interacción con un sistema tecnológico contemporáneo. La evaluación de la UX en entornos industriales se convierte en un desafío que merece especial atención, ya que la percepción de uso puede influir en la eficiencia y efectividad de la HRC.

Dentro de este contexto, analizar en profundidad los elementos fundamentales, explorando su estructura conceptual, taxonomía y modelo de evaluación se considera necesario. Este análisis, proporciona una visión comprensiva de cómo las interfaces y las interacciones en la HRI influyen en la dinámica colaborativa entre personas y robots.

Las interfaces unimodales se centran en una sola modalidad de comunicación, como puede ser la visual, la auditiva o la háptica. Esto puede simplificar la implementación y evitar posibles conflictos entre modalidades, lo que resulta en una UX más coherente y fácil de gestionar. Sin embargo, la limitación a una sola modalidad puede no aprovechar completamente la riqueza de la comunicación humana y restringir la capacidad de adaptación del sistema.

Por otro lado, las interfaces multimodales combinan varias modalidades de comunicación, lo que permite una representación más completa de la información y una mayor flexibilidad en la interacción. Esto puede ser especialmente beneficioso en situaciones complejas donde la información es diversa y requiere una representación más rica. Sin embargo, la implementación de interfaces multimodales puede ser más compleja y requerir un mayor esfuerzo de diseño y desarrollo.

La elección entre una interfaz unimodal y una multimodal depende en última instancia del contexto de uso y de las necesidades específicas de las personas usuarias. Es importante considerar factores como la naturaleza de la tarea, la capacidad perceptiva de las personas usuarias, la comodidad y la eficacia de la interacción. Una interfaz multimodal podría ser más adecuada cuando se requiere una representación rica de la información y la adaptación a diferentes preferencias de las personas usuarias, mientras que una interfaz unimodal podría ser más apropiada en situaciones donde la simplicidad y la coherencia son prioritarias.

Sin embargo, en cuanto a una interfaz multimodal, la combinación de las diferentes modalidades de la interfaz (visual, auditiva, háptica, etc.) para mejorar la comprensión y la percepción del usuario en la HRI aborda un desafío en el campo de la HRI. La integración efectiva de múltiples modalidades de *feedback* puede tener un impacto significativo en la calidad de la interacción entre personas y robots. La HRI se beneficiaría de una comunicación donde diferentes modalidades de *feedback* pueden transmitir información de manera más

efectiva y precisa. Sin embargo, la combinación óptima de estas modalidades implica una serie de consideraciones clave:

La complementariedad de las modalidades de interfaz es un aspecto fundamental que considerar. Cada modalidad de interfaz presenta sus propias ventajas y limitaciones. La combinación óptima debe aprovechar plenamente las fortalezas de cada tipo de interfaz y utilizarlas de manera complementaria para proporcionar información completa y coherente. Sin embargo, surge la pregunta de cuál podría ser esa combinación óptima. La comparación entre diferentes tipos de interfaces y cómo son percibidas por las personas usuarias es un campo de estudio que podría proporcionar información valiosa sobre cómo diseñar sistemas de HRI más efectivos y adaptados a las necesidades individuales de las personas usuarias. La búsqueda de esta combinación óptima constituye un paso relevante hacia la mejora de la UX en la HRI.

Además, es esencial determinar qué información es más relevante en un contexto particular. La priorización de la información puede guiar la asignación de modalidades de *feedback*. Dado que las preferencias de las personas usuarias pueden variar, la combinación de modalidades debe permitir la personalización. Las personas usuarias deben tener la capacidad de ajustar las modalidades de *feedback* según sus necesidades y preferencias individuales.

2.5.4.2 Impacto de las interfaces en los factores humanos y su evaluación

La HRI puede ser altamente demandante desde el punto de vista cognitivo, ya que implica procesar información de múltiples modalidades, tomar decisiones en tiempo real y adaptarse a situaciones cambiantes. Si no se gestionan adecuadamente, estas demandas cognitivas pueden generar estrés y fatiga en las personas, lo que a su vez puede afectar negativamente la percepción de la UX y la seguridad en el trabajo, y por ende a la aceptación del robot por parte de la persona. Por lo tanto, la investigación que aborda cómo minimizar esta carga cognitiva a través del diseño de interfaces y el *feedback* adecuado es de suma importancia para garantizar una HRI efectiva y segura.

La inclusión de soluciones adaptativas es un enfoque valioso, pero plantea desafíos significativos. La adaptabilidad es clave para atender las necesidades de las personas con diversas habilidades y discapacidades, pero la implementación efectiva de estas soluciones puede ser compleja. Además, la inclusión no se trata solo de abordar discapacidades físicas, sino también de considerar aspectos cognitivos y emocionales. Una pregunta crítica es cómo diseñar interfaces que no solo sean adaptables, sino también intuitivas y eficientes para todas las personas.

Tal y como se explicaba en el apartado 2.5.1.5, la confianza es fundamental para la colaboración efectiva entre personas y robots, pero es frágil y puede verse afectada por diversos factores, incluido el *feedback* proporcionado. Investigar cómo diseñar *feedback* que aumente la confianza del usuario es esencial, especialmente en entornos industriales donde la seguridad y la precisión son prioritarias.

Las métricas y medidas de la UX en la HRI plantean desafíos adicionales. La UX no se limita a aspectos funcionales, sino que también incluye dimensiones emocionales y subjetivas. Encontrar métricas adecuadas que reflejen estas dimensiones y que estén relacionadas con la percepción general de la experiencia es un área de investigación compleja pero necesaria.

2.5.4.3 La robótica afectiva

Dado que la UX es subjetiva y varía entre personas, es plausible que el *feedback* proporcionado en la HRI también deba adaptarse a las particularidades de cada persona. En este contexto, la robótica afectiva surge como una prometedora vía para ofrecer un *feedback* personalizado (Villani et al. 2018), ajustado al estado emocional de la persona usuaria y a las

acciones previstas del robot, determinando de manera precisa el canal, la frecuencia y la intensidad más apropiados (Cohen et al., 2018).

Además, se subraya que la interacción individualizada con el robot se define como una de las cinco categorías principales dentro del paradigma de la Industria 5.0 (Coronado et al., 2022). Este aspecto se erige como un pilar fundamental para lograr la interconexión y la sinergia entre las capacidades humanas y robóticas. Su papel radica en respaldar la calidad de la interacción y el compromiso en las relaciones a largo plazo entre personas y robots. Esta personalización contribuye a aumentar tanto la intención de uso como el uso efectivo de la tecnología robótica, a la vez que fomenta y mantiene la confianza de las personas usuarias en estos sistemas (Funk et al., 2020; Irfan et al., 2019).

Sin embargo, se plantea una interesante oportunidad de investigación: ¿cómo se puede medir el estado de ánimo de una persona en tiempo real durante la interacción con un robot? Los dispositivos fisiológicos se plantean como una herramienta útil y efectiva en este aspecto. No obstante, surge la cuestión de cuál sería la medida fisiológica óptima para lograr una evaluación precisa del estado emocional y que no resulte intrusiva. Además, es fundamental definir qué aspectos específicos se deben medir y cómo interpretar los datos obtenidos de estos dispositivos fisiológicos de manera adecuada, con el fin de permitir una adaptación efectiva de los robots y del *feedback* a las necesidades individuales de cada persona.

Estas cuestiones constituyen un terreno de investigación importante para mejorar la personalización y la eficacia de la HRI en entornos industriales. Sin embargo, antes de abordar estas preguntas, es imperativo comprender cómo las personas reaccionan ante los robots tanto en términos físicos como emocionales, identificar las posibles diferencias en la percepción entre individuos y desarrollar métodos adecuados para su medición.

Capítulo III

Modelo ITPX-ROBOTS

3 MODELO ITPX-ROBOTS

En el contexto de la HRI, el enfoque hacia la optimización de la UX ha emergido como un pilar fundamental en la confluencia de la tecnología robótica y las interacciones humanas. El diseño y desarrollo de sistemas robóticos efectivos y colaborativos trasciende el ámbito de la mera eficiencia técnica, abarcando la calidad y la satisfacción intrínseca de la persona usuaria en la relación persona-robot.

Este capítulo, por tanto, se adentra en la concepción y presentación de un modelo conceptual estructurado que articula la naturaleza compleja y multidimensional de la HRI y su evaluación.

3.1 Fundamentos teóricos

En el proceso de desarrollo de un modelo de evaluación de la Experiencia de Interacción Persona-Robot (HRIX, por sus siglas en inglés) en entornos industriales, resulta esencial comprender los conceptos fundamentales que subyacen en la HRI y en la percepción de la persona usuaria. En esta sección, se presentan los elementos teóricos esenciales que dan forma a este modelo.

Para la creación de un nuevo modelo, se han tenido en cuenta diferentes marcos teóricos. Entre ellos, se ha identificado la falta de un método de evaluación estructurado que se adapte a diferentes contextos de HRI y diferentes etapas del proceso de diseño. El marco propuesto por Prati et al. (2021), denominado ciclo UX en HRI, representa el enfoque más cercano, sin embargo, muestra carencias en la fase de evaluación. En contraposición, el modelo de evaluación ITPX (Mazmela, 2020), se presenta como una solución integral, abordando tanto la UX como la aceptación tecnológica al evaluar HMIs industriales. En este sentido, el modelo ITPX (Mazmela, 2020) destaca que, para lograr una evaluación exhaustiva de la interacción, es esencial considerar el ajuste entre las características de la persona, las características del robot y las características del contexto.

Partiendo del marco de referencia proporcionado por el modelo ITPX (Mazmela, 2020), se presenta un nuevo enfoque que propone un procedimiento de análisis fundamentado en tres componentes principales. En primer lugar, se utiliza un indicador que se evalúa mediante una lista de comprobación diseñada específicamente para evaluar la usabilidad de los sistemas de HRI mediante el uso de heurísticas. En segundo lugar, se realiza un análisis cuantitativo de la ejecución de tareas, en el cual se recopilan métricas de rendimiento, como el tiempo necesario para completar las actividades y un índice de eficiencia. Por último, se incorpora la perspectiva de las personas operarias al sistema, evaluando su percepción y experiencia en la interacción con el sistema. La innovación clave de este enfoque radica en la integración de estos tres enfoques, marcando una diferencia significativa con respecto a los estudios existentes.

A continuación, se exponen las características que el nuevo modelo debe integrar:

- i) **Enfoque holístico.** El modelo debe integrar un modelo holístico, de manera que recoja las percepciones de las personas en todo momento, es decir, antes, durante y después de la interacción. Además, será necesario que el nuevo modelo incluya diferentes métodos de medición en diferentes momentos de la interacción, cuantitativos y cualitativos, objetivos y subjetivos, con el fin de garantizar la interpretación de los datos y el entendimiento de todo el flujo de la interacción. La triangulación que involucra medidas fisiológicas, perceptuales y de

desempeño para la evaluación de la UX proporcionan información útil que puede mejorar la experiencia del robot.

- ii) **La perspectiva de la UX.** De acuerdo con la perspectiva de la UX, las percepciones y respuestas de una persona que resultan del uso o el uso anticipado de un producto, sistema o servicio, incluidas las emociones, creencias, preferencias, percepciones, respuestas físicas y psicológicas, comportamientos y logros de las personas usuarias que se producen antes, durante y después del uso. En este sentido, la UX tiene un carácter temporal, es decir, incluye las percepciones de las personas antes, durante y después de la interacción. Sin embargo, no se ha identificado ningún estudio experimental que incluya este carácter temporal a la hora de realizar la evaluación. Además, el carácter experiencial resulta de gran importancia y, por lo tanto, se han de considerar los tres componentes clave: usabilidad, emociones y apariencia.
- iii) **Evaluación fisiológica.** La emoción humana, el comportamiento y la percepción al utilizar un sistema a menudo están regulados por sus respuestas neurológicas y fisiológicas. Las herramientas fisiológicas permiten realizar una evaluación más objetiva, ya que proporcionan información sin sesgo retrospectivo. En la literatura, tan solo se ha identificado un estudio que utilice la monitorización fisiológica con robots (Harriott et al., 2013) que utiliza un dispositivo ECG, y tan solo dos que lo utilicen en el ámbito de la Industria 4.0 (Peruzzini et al., 2018, 2020), que utilizan *wearables* que monitorizan la frecuencia respiratoria, PPG, HR y ST. Ante esta carencia, el nuevo modelo deberá integrar la monitorización de la persona usuaria mediante herramientas fisiológicas.
- iv) **Variables cinemáticas.** Se han de tener en cuenta las variables cinemáticas que están relacionadas con las características del robot, como pueden ser el ruido o la velocidad, con el fin de conocer el efecto de estos en la percepción de las personas y en su desempeño.
- v) **Aspectos subjetivos.** Se han de tener en cuenta las emociones como la fidelidad, la estimulación y la confianza, e incluso tener en cuenta la apariencia, puesto que, en la literatura, se ha prestado poca atención a los efectos emocionales de las personas usuarias con respecto a la impresión estética. Resultaría interesante complementar el estudio con herramientas tradicionales como los cuestionarios o las entrevistas, que permiten conocer la percepción de las personas usuarias de forma subjetiva.
- vi) **Seguridad y riesgo percibido.** En el contexto de interacción con un robot, la seguridad y el riesgo percibido es un factor determinante, puesto que impacta directamente en el desempeño y en las emociones de la persona. De hecho, el marco teórico propuesto por Villani et al. (2018) pone en el centro la seguridad, y el marco propuesto por Meissner et al. (2020) indica que uno de los factores primarios influyentes es el riesgo percibido.

La combinación de las diferentes herramientas permitirá una evaluación más completa de la UX, es decir, una evaluación que incluya las medidas fisiológicas y las obtenidas mediante percepciones subjetivas será la óptima para este contexto.

3.1.1 Marco conceptual HRIX

En base a las oportunidades identificadas, se propone el desarrollo de un nuevo marco conceptual denominado *Human-Robot Interaction Experience* (HRIX), que se muestra en la Figura 24. Partiendo del modelo CUE propuesto por Thüring & Mahlke (2007), que integra aspectos experienciales e instrumentales, el presente estudio propone el siguiente marco conceptual. Según Thüring & Mahlke (2007), este marco aborda las características del sistema (en este caso, el robot), la persona y el contexto. El marco conceptual HRIX sostiene que la

interacción entre un robot y una persona viene determinada por una combinación de las características del robot, las características de la persona usuaria y los factores contextuales.

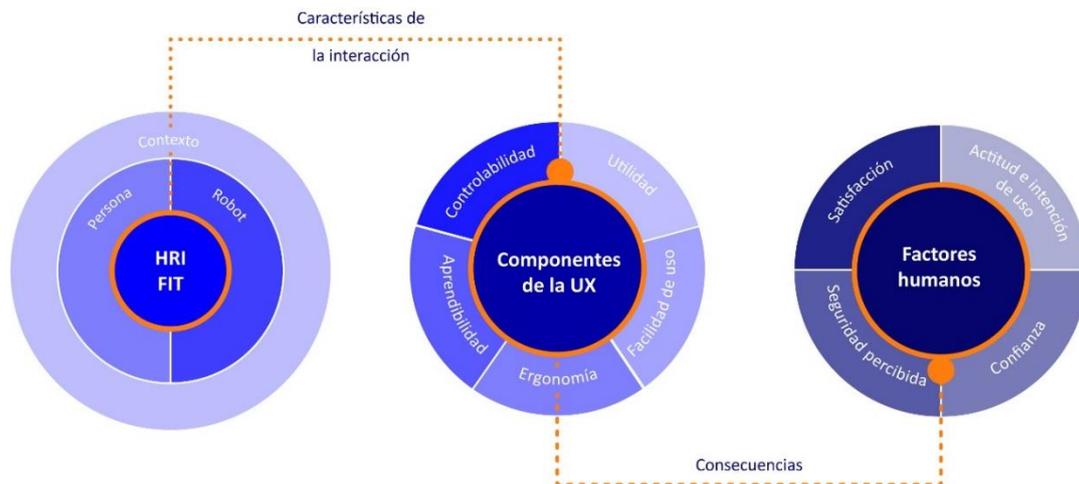


Figura 24: Marco conceptual HRIX.

El ajuste entre las características de la interacción, a partir de ahora “HRI FIT” (Figura 30), tiene un impacto en los componentes de la UX, que en el marco conceptual HRIX (Figura 24) se definen:

- La **utilidad** se relaciona con el grado en que una persona percibe que el uso de un robot específico mejoraría su desempeño laboral. Implica la evaluación de la eficacia del sistema en el contexto de tareas laborales específicas.
- La **facilidad de uso** se refiere al grado en que una persona cree que el uso de un robot en particular sería libre de esfuerzo y fácil de aprender. Esta dimensión se enfoca en la accesibilidad y la simplicidad de la interacción con el robot.
- La **ergonomía**, refiriéndose, por un lado, a cómo el diseño del robot se adapta a las características anatómicas, antropométricas, fisiológicas y biomecánicas de las personas usuarias en relación con la actividad física (International Ergonomics Association, 2019); y, por otro lado, al grado en que el uso del robot se ajusta a las capacidades cognitivas de las personas usuarias, lo que incluye la percepción humana, el procesamiento mental y la memoria.
- La **aprendibilidad** significa el grado en que las personas usuarias pueden realizar con éxito una tarea cuando se encuentran con robot por primera vez y el grado en que pueden llegar a ser competentes en esa tarea con repeticiones de uso cada vez mayores (Joyce, 2019).
- La **controlabilidad** se refiere al grado en que una persona se siente en control de una tecnología y sus acciones, y la capacidad de modificar su comportamiento de acuerdo con sus necesidades o preferencias (Sarker & Wells, 2003).

Por lo tanto, si el ajuste entre las características de interacción es bueno, los componentes de la UX pueden conducir a la mejora de la experiencia global, lo que se traduce en consecuencias sobre los factores humanos, que en este contexto se consideran (Figura 24):

- **Actitud e intención de uso:** La actitud hacia la tecnología y la intención de uso se refieren a la evaluación general positiva o negativa que tiene una persona hacia un sistema y su disposición a utilizarlo en el futuro (Bröhl et al., 2019). Estos factores pueden influir en la adopción y el compromiso continuo con la tecnología.

- **Confianza:** La confianza implica el grado en que una persona se siente segura y cómoda al utilizar un sistema. Una mayor confianza puede mejorar la eficacia de la interacción y la percepción general de la tecnología (Hancock et al., 2011).
- **Seguridad percibida:** La seguridad percibida se relaciona con la percepción de una persona sobre la seguridad al interactuar con un sistema. Implica evaluar si el sistema es seguro de usar y si se pueden evitar posibles riesgos o problemas (Akalín et al., 2019).
- **Satisfacción:** La satisfacción refleja el grado en que una persona está contenta con la experiencia general de usuario al utilizar un sistema. Una mayor satisfacción está asociada con una experiencia general más positiva (Paliga, 2023).

3.1.2 Factores influyentes y métricas en la HRI

En contextos de HRI en entornos industriales, la persona, el robot y la interacción entre ambos también se ven influenciados por el contexto. Como se refleja en el TAM (Rogers et al., 2014) y tal y como lo reflejan Meissner et al. (2020), es necesario que exista un ajuste entre las características de la persona, las características del robot y las características del contexto para que la interacción sea fluida y satisfactoria.

En este sentido, se ha desarrollado la taxonomía *Human-Robot-Context* (HUBOXT), que se compone de las características de la persona, las características del robot y las características del contexto. A su vez, las características del contexto se dividen en niveles macro y micro, siguiendo el enfoque del Modelo de Experiencia de Usuario Unificada desarrollado por von Saucken & Gomez (2014), que busca abordar el aspecto complejo, dinámico y evolutivo de la experiencia a lo largo del tiempo (Figura 25).

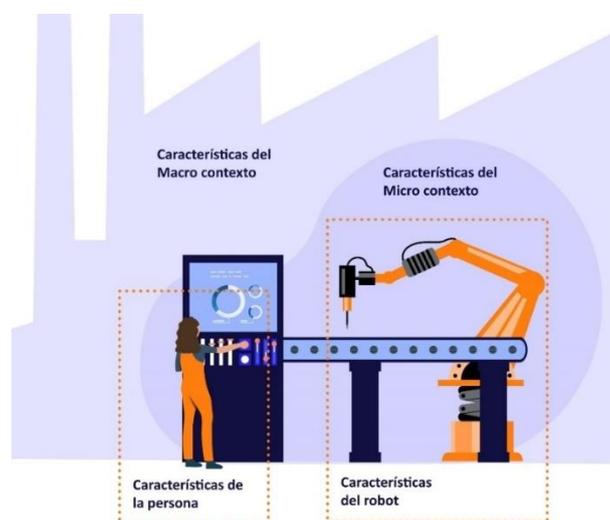


Figura 25: Visión general de la taxonomía HUBOXT.

La taxonomía HUBOXT, que se muestra en la Figura 26, constituye un enfoque integral para entender cómo diversas características relacionadas con la persona, el robot y el contexto se entrelazan y afectan la percepción y la experiencia del usuario en interacciones humano-robot. Esta taxonomía se basa en la premisa de que una interacción efectiva entre personas y robots va más allá de solo considerar aspectos técnicos o de diseño, ya que también involucra elementos intrínsecos a la persona y su entorno. A continuación, se describen en detalle las características clave de la taxonomía HUBOXT:

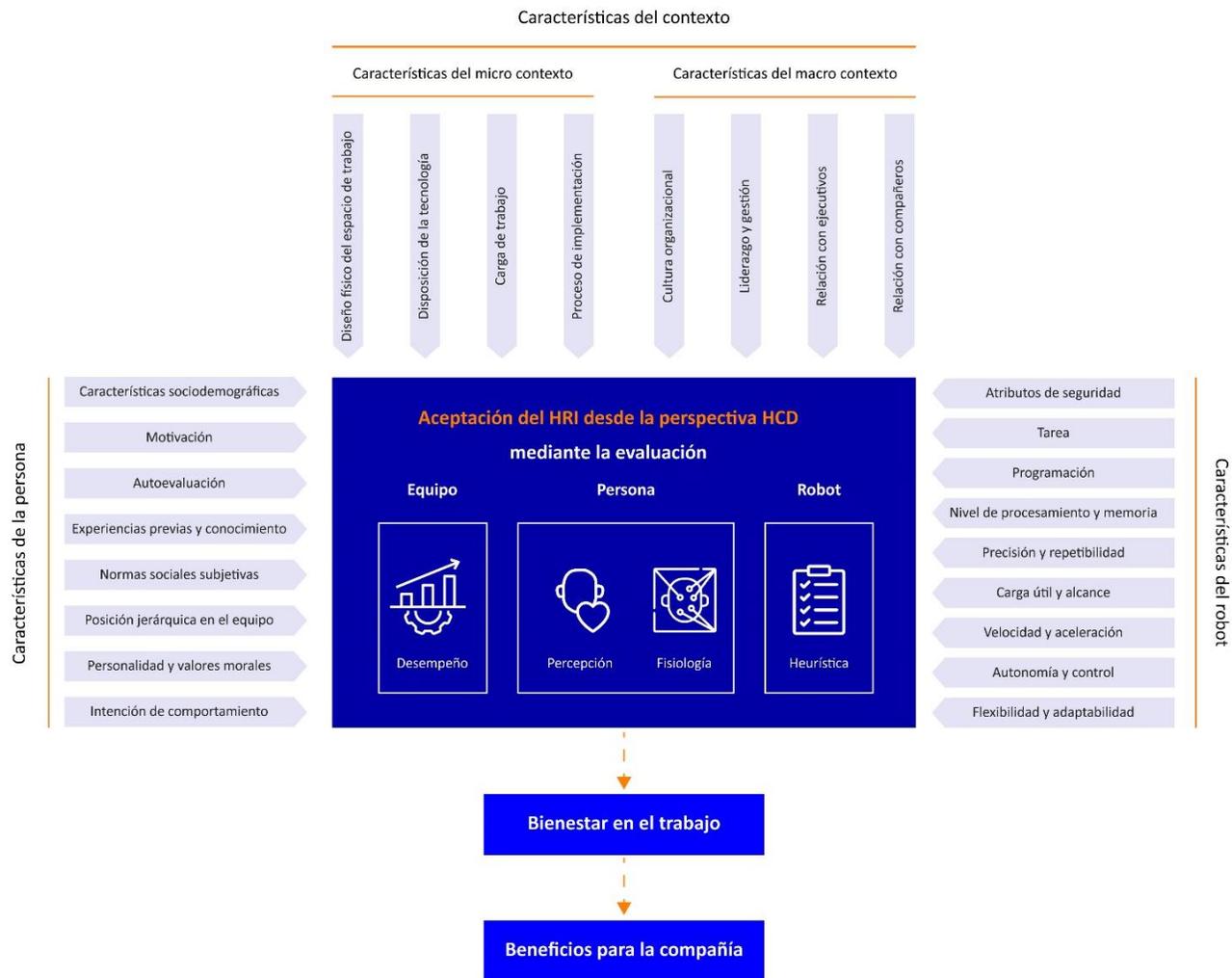


Figura 26: Modelo conceptual de la taxonomía HUBOXT para elaborar el modelo de evaluación.

3.1.2.1 Características de la persona

En la Figura 27 se exponen las características individuales que ejercen influencia sobre la HRI, las cuales se detallan de manera exhaustiva a continuación:



Figura 27: Características de la persona que influyen en la aceptación de la HRI.

En lo que corresponde a las características individuales de la persona desempeñan un papel crucial en la forma en que se establece dicha interacción. En particular, **las características sociodemográficas**, como el género, la edad y el nivel de educación, influyen en las expectativas, habilidades, normas sociales y preferencias de las personas con respecto a los robots. En el caso del género, existen estudios que sugieren que hombres y mujeres pueden tener diferentes expectativas y preferencias en cuanto a las características del robot y su uso. Sin embargo, es preciso seguir investigando en las diferencias de género a la hora de interactuar con un robot. Es importante tener en cuenta que las generalizaciones no son aplicables en todos los casos y que la influencia del género y la edad en la HRI puede variar según el contexto y las experiencias individuales de cada persona. Por tanto, se debe tener un enfoque individualizado y considerar las características personales de cada usuario al diseñar y desarrollar robots para una interacción efectiva y satisfactoria.

La **motivación** es otro factor importante que influye en la HRI y en la aceptación de la tecnología. Las personas motivadas tienden a estar más dispuestas a interactuar con los robots, a explorar sus características y funcionalidades, y por tanto adaptarse a las nuevas formas de interacción (Welfare et al., 2019). Además, suelen ser más tolerantes ante los errores y limitaciones del robot. Por el contrario, la falta de motivación puede disminuir la disposición de las personas a utilizar robots y aceptarlos para su trabajo en el día a día. En el ámbito laboral, la motivación de las personas participantes puede variar entre motivación intrínseca y extrínseca. Algunos individuos consideran su trabajo como un medio para ganarse la vida, mientras que otros encuentran satisfacción personal en su labor (Meissner et al., 2020). Aquellos con motivación intrínseca tienden a buscar variedad y desafíos en sus tareas, mostrando una mayor apertura a la adopción de cambios, como la HRI.

Por otro lado, la **autoevaluación**, que se refiere a las creencias de control individuales en relación con su capacidad para llevar a cabo una tarea (Venkatesh & Bala, 2008), es un concepto relevante que influye en la percepción de las personas trabajadoras sobre su

habilidad para adaptarse a nuevas situaciones, como podría ser la HRI. En su mayoría, esta noción se aborda mediante el concepto de autoeficacia (Meissner et al., 2020).

Las **experiencias previas y el conocimiento** también tienen una influencia significativa en la HRI y la aceptación tecnológica. Aunque las experiencias positivas pueden mejorar la aceptación y uso de los robots, las experiencias negativas, como la incapacidad de interactuar con la tecnología debido a un deterioro de habilidades, pueden tener el efecto opuesto. Además, las personas participantes con un alto nivel de conocimiento tecnológico tienden a tener menos temores irracionales, como la creencia de que los robots son más inteligentes que las personas (Welfare et al., 2019). Sin embargo, también pueden expresar preocupaciones específicas, como posibles fallas técnicas. Por otro lado, la experiencia práctica derivada del trabajo con robots parece fortalecer la confianza de las personas en sus habilidades para interactuar con ellos, lo que puede influir en su actitud hacia la HRC.

Las **normas sociales subjetivas** emergen como un componente esencial que puede modelar significativamente la aceptación y adopción de nuevas tecnologías, incluidos los sistemas de HRI. Estas normas reflejan las creencias y opiniones que los individuos mantienen acerca de la percepción social y la aceptabilidad de interactuar con tecnologías específicas. La influencia de las normas sociales subjetivas se manifiesta en la percepción de las personas sobre la adecuación y aceptabilidad de la interacción con robots. Estas normas pueden ser determinantes en la formación de actitudes y comportamientos hacia la adopción de tecnologías, ya que los individuos tienden a alinear sus decisiones con las expectativas sociales percibidas. Por lo tanto, las normas sociales subjetivas pueden actuar como facilitadoras o barreras para la adopción de los sistemas robóticos. Identificar cómo las creencias compartidas y las expectativas sociales afectan las actitudes individuales hacia los robots puede proporcionar información valiosa para el diseño de estrategias que promuevan una mayor aceptación y adopción de estas tecnologías.

La **posición jerárquica en el equipo** podría tener un impacto significativo en la actitud de las personas trabajadoras hacia el robot. De acuerdo con Meissner et al. (2020), las personas trabajadoras en posiciones de liderazgo o con responsabilidades adicionales más allá de las tareas de ensamblaje tienden a ver al robot como una oportunidad para familiarizarse con una tecnología innovadora. Por otro lado, aquellos responsables únicamente del ensamblaje podrían sentirse amenazados por la introducción de un robot, ya que esto podría llevar a sentir preocupación e incertidumbre por una posible pérdida de empleo y de seguridad ocupacional. Estas perspectivas divergentes resaltan la importancia de considerar los roles y responsabilidades de las personas trabajadoras al implementar sistemas HRC en el lugar de trabajo.

La **personalidad y los valores morales** también influyen en la aceptación de la tecnología (Devaraj et al., 2008; Svendsen et al., 2013). Rasgos de personalidad como el optimismo, la apertura a nuevas experiencias y la extraversión parecen estar asociados con una actitud positiva hacia la tecnología. En la HRI, estudios han identificado una relación positiva entre la extraversión y las actitudes hacia los robots (Damholdt et al., 2015; Müller & Richert, 2018; Reich & Eysel, 2013). De acuerdo con Meissner et al. (2020), lo mismo ocurre con los valores morales. Las personas participantes indican diferentes posturas éticas hacia los robots y la HRI en general, lo que parece estar relacionado con sus actitudes hacia la implementación de la HRI en la empresa. Sin embargo, esto parece ser especialmente cierto para aquellas personas que sienten que no se ven personalmente afectadas por la introducción de la HRI, como los líderes de equipo, por ejemplo. Ellos tienden a preocuparse más por los efectos sociales generales de la introducción de robots en el lugar de trabajo, mientras que las personas que temen ser reemplazadas se preocupan más por su propia situación.

Por último, la **intención de comportamiento** emerge como un factor fundamental con una influencia significativa en la aceptación y adopción de nuevas tecnologías, incluyendo sistemas de HRI. Esta intención se refiere a la disposición intrínseca de un individuo para llevar a cabo un comportamiento específico, como la interacción con un robot. Puede ser considerada como un precursor directo de la acción, representando la etapa crucial entre las actitudes y las conductas efectivas. En este sentido, la disposición de una persona para interactuar con un robot puede ser moldeada por diversas variables contextuales y personales, y todos los factores previamente mencionados influyen.

3.1.2.2 Características del robot

En la Figura 28 se exponen las características del robot que ejercen influencia sobre la HRI, las cuales se detallan de manera exhaustiva a continuación.



Figura 28: Características del robot que influyen en la aceptación de la HRI.

Los atributos de seguridad desempeñan un papel fundamental en la aceptación de un robot en entornos industriales. Estos atributos representan un requisito mínimo esencial que debe cumplir cualquier robot destinado a interactuar y colaborar con personas. Están relacionados con las características intrínsecas del robot diseñadas para proteger la integridad física y el bienestar de las personas que interactúan con él. Los atributos de seguridad de un robot comienzan con la implementación de medidas de seguridad sólidas (Marvel et al., 2020; Villani et al., 2018). Estas medidas pueden incluir sensores y sistemas de detección avanzados que permiten al robot detectar la presencia de personas en su entorno y ajustar su comportamiento en consecuencia. También podrían implicar sistemas de parada de emergencia que permiten a las personas trabajadoras detener rápidamente las operaciones del robot si se detecta una situación peligrosa. Además de las características técnicas, la seguridad también se relaciona con la programación y el diseño del robot. Una programación diseñada para evitar movimientos bruscos o inesperados que puedan representar un riesgo para las personas trabajadoras. Es importante destacar que los atributos de seguridad no solo influyen en la seguridad física, sino también en la percepción de seguridad por parte de las personas usuarias, y por ende en la confianza (Tsai et al., 2014).

La tarea alude a las acciones específicas que el robot se encuentra encargado de ejecutar, abarcando una gama diversificada de funciones que engloban desde la manipulación de objetos hasta el ensamblaje de componentes, la realización de procesos de soldadura, el

pintado y el transporte de materiales, entre otras múltiples aplicaciones industriales. En este sentido, la tarea ejerce un impacto multifacético, determinante para la definición del diseño del robot en función de las necesidades específicas. La naturaleza de la tarea incide en la tipología de robot requerida, la capacidad de carga que debe ostentar, la velocidad intrínseca de operación, así como en los niveles de precisión y repetibilidad demandados (Colceriu et al., 2022; Pardo et al., 2016; Salem et al., 2015). Asimismo, la complejidad inherente a la programación es un aspecto íntimamente vinculado a la tarea, ya que esta última puede variar sustancialmente en su grado de intrincación y demandar enfoques programáticos igualmente variados.

La programación del robot, como componente esencial en su funcionalidad y comportamiento, impacta en la forma en que el robot interactúa y desempeña sus tareas, lo que puede determinar que las personas usuarias perciban al robot como una herramienta útil y adecuada para sus necesidades laborales. Además, la adaptabilidad y flexibilidad de la programación son factores cruciales que impactan en la aceptación del robot. Una programación que permite al robot ajustarse a diferentes situaciones y contextos contribuye a que las personas usuarias experimenten una mayor comodidad y eficacia en su interacción con el robot. Por otro lado, una programación rígida y limitada puede generar frustración y rechazo, ya que el robot podría no satisfacer plenamente las demandas y expectativas de las personas usuarias en diversas circunstancias. Además, la programación puede influir en la percepción de seguridad de las personas usuarias hacia el robot (Colceriu et al., 2022; Salem et al., 2015). Un robot programado de manera coherente y predecible puede generar confianza en las personas usuarias, ya que estos pueden anticipar su comportamiento y resultados. En contraste, una programación errática o poco confiable puede generar desconfianza y temor, lo que podría impactar negativamente en la aceptación del robot. Asimismo, la programación puede afectar la percepción de competencia y utilidad del robot. Un robot que se desempeña de manera eficiente y precisa en sus tareas gracias a una programación adecuada es más propenso a ser valorado positivamente por las personas usuarias. Por el contrario, un robot con dificultades para ejecutar tareas debido a problemas de programación puede llevar a una percepción de inutilidad y rechazo.

En lo que corresponde a las características del robot, los robots industriales requieren un alto **nivel de procesamiento y memoria** para realizar tareas complejas en tiempo real. El procesamiento se refiere a la capacidad del robot para ejecutar tareas complejas de manera rápida y eficiente. La memoria, por su parte, se refiere a la capacidad del robot para almacenar información necesaria para la realización de las tareas. En un entorno industrial, los robots pueden realizar tareas que requieren una gran cantidad de cálculos y decisiones en tiempo real, como la identificación de piezas o la toma de decisiones en función de la información sensorial que reciben (Marvel et al., 2020; Negi, 2021). Para poder realizar estas tareas, los robots industriales necesitan un alto nivel de procesamiento y memoria.

Además, **la precisión** se refiere a la capacidad del robot para ejecutar una tarea con una alta exactitud en cuanto a la posición, velocidad y aceleración. La precisión es importante en aplicaciones donde se requiere la manipulación de objetos pequeños o en trabajos que necesitan una alta precisión, como en la fabricación de piezas de alta tecnología. **La repetibilidad**, por otro lado, se refiere a la capacidad del robot para ejecutar una tarea de manera constante en diferentes ciclos de trabajo. La repetibilidad es importante en aplicaciones donde se requiere una alta precisión a lo largo de múltiples ciclos de trabajo, como en la automatización de procesos de producción en serie. La precisión y la repetibilidad están estrechamente relacionadas y deben ser evaluadas juntas al seleccionar un robot. Una alta precisión no garantiza una alta repetibilidad, y viceversa. La precisión puede ser influenciada por factores externos, como las condiciones ambientales, mientras que la repetibilidad puede verse afectada por el desgaste de los componentes mecánicos del robot.

Es importante tener en cuenta que la precisión y repetibilidad del robot pueden variar según la tarea que realiza (Bhalaji et al., 2021). En algunos casos, es posible que un robot sea muy preciso, pero no sea repetible en una tarea determinada, mientras que en otras tareas puede tener una alta repetibilidad, pero no una precisión extrema.

La carga útil se refiere a la cantidad máxima de peso que el robot puede manipular. **El alcance**, por otro lado, se refiere a la distancia máxima que el robot puede mover su extremo efecto. La percepción de que el robot puede desempeñar tareas que requieren fuerza física y manejo de objetos pesados puede aumentar la aceptación al reconocer la contribución y complementariedad del robot en el entorno de trabajo.

La velocidad y la aceleración pueden afectar directamente el tiempo necesario para completar una tarea, así como la precisión y repetibilidad de las acciones del robot. La velocidad y aceleración también pueden influir en la seguridad del robot y las personas que trabajan cerca de él. Si un robot se mueve demasiado rápido o tiene una aceleración demasiado alta, puede ser más difícil para las personas predecir su movimiento y evitar colisiones potenciales. Es importante encontrar un equilibrio entre la velocidad y la aceleración adecuadas para la tarea que realiza el robot y las condiciones en las que opera (Arai et al., 2010; Dehais et al., 2011; Dragan et al., 2015; Hanajima et al., 2017; Lasota & Shah, 2015).

La autonomía, definida como la capacidad del robot para tomar decisiones y ejecutar tareas de manera independiente, **y el control**, que se refiere a la habilidad de la persona usuaria humano para dirigir las acciones y el comportamiento del robot, se entrelazan en un equilibrio que tiene importantes consecuencias para la percepción y aceptación del robot en el entorno laboral (Gervasi, Digiario, et al., 2020; Tsarouchi et al., 2016). De acuerdo con Meissner et al. (2020), la autonomía del robot puede generar una variedad de reacciones emocionales y cognitivas en las personas trabajadoras. Una mayor autonomía puede ser percibida como una característica positiva, especialmente cuando permite al robot asumir tareas repetitivas, peligrosas o tediosas, liberando a las personas trabajadoras de actividades monótonas y de riesgo. Sin embargo, un nivel excesivo de autonomía puede también generar preocupaciones sobre la pérdida de control humano sobre las operaciones y la incertidumbre en relación con las decisiones del robot. La sensación de estar relegado a un papel pasivo, donde la persona trabajadora se convierte en espectador de las acciones del robot, puede desencadenar sentimientos de desconfianza y resistencia hacia la tecnología. El control humano sobre el robot se vincula con la capacidad de las personas trabajadoras para influir en el comportamiento y las acciones del robot según las necesidades operativas y las preferencias individuales. La posibilidad de dirigir al robot y proporcionar instrucciones específicas puede generar un sentido de empoderamiento y participación activa en el proceso de producción, mejorando la sensación de control y la confianza en la tecnología. La capacidad de ajustar el comportamiento del robot a situaciones cambiantes y requerimientos específicos puede aumentar la percepción de utilidad y adaptabilidad de la tecnología en el entorno laboral.

Por último, la **flexibilidad y la adaptabilidad** del robot son componentes fundamentales para su aceptación. La adaptabilidad a las personas trabajadoras constituye un factor determinante en la percepción positiva y la aceptación de la HRI. Las personas participantes expresan la expectativa de que los robots sean ajustables a las demandas físicas de las personas, como la altura del cuerpo, así como a los distintos modos individuales de trabajo, como el ritmo de trabajo. Esta adaptabilidad es fundamental para garantizar alivio tanto físico como mental a las personas trabajadoras. Además, es importante que los robots sean capaces de asumir diversas tareas en la producción, lo que subraya la necesidad de flexibilidad en los procesos industriales. Sin embargo, esta flexibilidad debe ser percibida como beneficiosa y estratégica para la eficiencia de la producción (Negi, 2021). Es fundamental que los robots proporcionen ventajas claras sobre los métodos manuales, lo que justifica su implementación en el proceso de ensamblaje.

3.1.2.3 Características del contexto

Por último, el entorno industrial y laboral también influyen en la aceptación del robot por parte de las personas operarias. En este sentido y de acuerdo con von Saucken & Gomez (2014) se diferencian el micro contexto, que hace referencia al ámbito del propio espacio de trabajo, y el macro contexto, que abarca de manera holística todos los aspectos relacionados con la organización y la empresa en su conjunto. Sin embargo, es importante resaltar que algunos factores tienen elementos tanto de micro como de macro contexto, ya que están relacionados con la interacción entre las personas trabajadoras y la tecnología a nivel individual (micro) y con la dinámica organizativa y cultural más amplia (macro). Los factores influyentes en la adopción del robot (Figura 29), interactúan de manera compleja, lo que puede hacer que la distinción entre el micro contexto y macro contexto sea sutil y tenga un impacto en la aceptación de la HRI.

Características del contexto

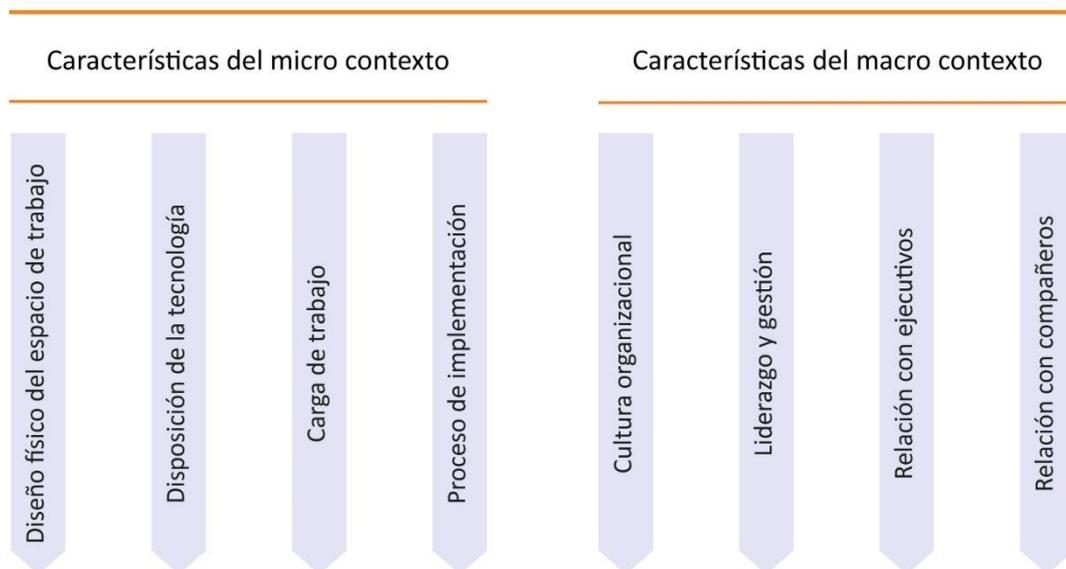


Figura 29: Características del contexto que influyen en la aceptación de la HRI.

3.1.2.3.1 Características del Micro contexto

El **diseño físico del espacio de trabajo** en un entorno industrial desempeña un papel crucial en la aceptación de un robot por parte de las personas operarias. La disposición del espacio, la ergonomía, la iluminación y la ubicación de los equipos pueden influir significativamente en cómo las personas trabajadoras interactúan y se sienten al trabajar con robots en su entorno laboral. El diseño físico del espacio en un contexto industrial puede influir en la aceptación de un robot al afectar la comodidad, la seguridad y la eficiencia de la interacción entre personas y robots (Svarny et al., 2019). Un diseño que promueva una colaboración fluida, ergonomía adecuada, iluminación favorable y una disposición armoniosa de equipos puede contribuir en gran medida a que las personas operarias se sientan cómodas y positivas al trabajar con robots, lo que a su vez mejora la aceptación y efectividad de la tecnología en el lugar de trabajo.

La **disposición de la tecnología** implica la ubicación planificada y la integración de robots en las operaciones existentes en un contexto industrial (Svarny et al., 2019). Si los robots se ubican de manera que faciliten su interacción con las personas trabajadoras y mejoren la eficiencia de las tareas, es más probable que las personas operarias los vean como herramientas útiles y valiosas. La integración de los robots con las operaciones existentes es igualmente crucial. Si los robots complementan y mejoran las tareas realizadas por las

personas trabajadoras, es más probable que estos últimos vean la tecnología como una ayuda en lugar de una amenaza.

La **carga de trabajo** en términos de cantidad y complejidad de tareas realizadas en colaboración con los robots en un contexto industrial influye directamente en la aceptación de la tecnología por parte de las personas operarias. Una carga de trabajo equilibrada y bien diseñada, que aproveche las fortalezas de las personas operarias humanas y los robots, puede aumentar la percepción positiva de la tecnología y, en última instancia, promover su aceptación en el entorno laboral. Por otro lado, una carga de trabajo excesiva o mal diseñada puede generar resistencia y obstáculos en la adopción exitosa de los robots (Prewett et al., 2010).

3.1.2.3.2 Características del Macro contexto

La **cultura organizacional** se refiere al conjunto de valores, normas, creencias y actitudes compartidas dentro de una empresa. Esta cultura influye en la forma en que los empleados perciben y se relacionan con la tecnología y la automatización, lo cual a su vez afecta su disposición a aceptar y trabajar en colaboración con los robots (Meissner et al., 2020).

El **liderazgo y la gestión** ejercen una influencia directa en la aceptación de los robots en un contexto industrial. Un enfoque de liderazgo participativo, un apoyo gerencial sólido y una toma de decisiones inclusiva pueden crear un entorno propicio para que las personas trabajadoras acepten y colaboren con los robots de manera efectiva.

La **relación con los ejecutivos** y el estilo de gestión en un contexto industrial tienen un impacto significativo en la aceptación de los robots por parte de las personas trabajadoras. La confianza en los ejecutivos, la comunicación efectiva, la transparencia en la toma de decisiones y el involucramiento de los empleados son aspectos cruciales para promover una actitud positiva y una mayor disposición a colaborar con los robots en el entorno laboral.

La **relación con los compañeros** y el ambiente de trabajo colaborativo juegan un rol crucial en la aceptación de la HRI en el entorno laboral. La solidaridad y el respaldo emocional proporcionados por un equipo cooperativo ayudan a las personas trabajadoras a superar sus inseguridades iniciales y a cultivar una actitud positiva. La presencia de un equipo comprometido y positivo puede amortiguar los efectos negativos de la incertidumbre y promover una adopción exitosa de la tecnología.

3.1.2.4 *El ajuste entre las características de la persona, el robot y el contexto*

El ajuste entre las características de la interacción determina si la interacción es satisfactoria, fluida y eficaz. Para lograr una interacción óptima, debe haber un ajuste entre las características de la persona, las características del robot y las características contextuales (Figura 30). Que haya un ajuste significa que el diseño y las capacidades del robot deben coincidir con las características y necesidades de las personas trabajadoras, y también deben tenerse en cuenta los factores contextuales, como el entorno de trabajo. Un buen ajuste entre estos factores dará como resultado una experiencia positiva para la persona usuaria y una mayor aceptación del sistema.

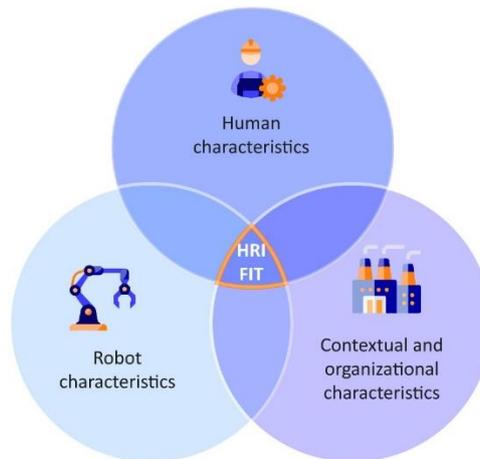


Figura 30: HRI FIT - El ajuste entre las características de la persona, las características del robot y las características del contexto.

3.1.3 De los factores influyentes a la evaluación

Los factores influyentes previamente delineados tienen un impacto directo en el grado de aceptación del robot y en la calidad de la UX. La evaluación de estos aspectos se concentra en los actores principales de la HRI, es decir, la persona y el robot, y en menor medida las características específicas del micro contexto. Es importante señalar que, en este enfoque y modelo, se han excluido las características del macro contexto, las cuales se relacionan principalmente con aspectos organizativos y empresariales (Figura 31).



Figura 31: Marco teórico de la evaluación.

Para evaluar el ajuste en HRI, se proponen métricas objetivas que permitan medir de manera precisa cómo se desenvuelven la persona y el robot en diversas tareas y situaciones. En este sentido, se propone emplear indicadores de desempeño, que ofrecen una visión cuantitativa y medible de la actuación de la persona y el robot trabajando juntos. Coronado et al. (2022) proporcionan una serie de métricas de desempeño útiles, como el tiempo de ejecución de las tareas o el índice de producción, entre otros.

Desde la perspectiva HCD, no solo es importante medir el desempeño objetivo del equipo persona-robot, sino que también se debe considerar la experiencia subjetiva de la

persona que interactúa con el robot. Esta evaluación subjetiva debe abarcar tanto aspectos hedonómicos como pragmáticos para lograr una comprensión completa de la calidad de la interacción y la satisfacción de la persona usuaria. Además de la evaluación subjetiva, para enriquecer el modelo con una visión objetiva de la persona, es importante agregar aspectos fisiológicos. Estos agregan una dimensión adicional a la comprensión de la experiencia ya que pueden revelar respuestas emocionales que no siempre son evidentes a través de la evaluación subjetiva, y, además, está libre de sesgos. La combinación de ambas perspectivas ofrece una comprensión completa de la persona usuaria.

Por último, la evaluación del robot y del micro contexto en el que este opera puede llevarse a cabo mediante la aplicación de un listado de heurísticas. Este enfoque se convierte en un componente esencial para analizar y mejorar la interacción entre el robot y su espacio de trabajo en el que se encuentra. La utilización de un conjunto de heurísticas para evaluar tanto el robot como su micro contexto proporciona una estructura sistemática para identificar y abordar problemas en la HRI. Esto ayuda a evaluar si el robot como su entorno están diseñados de manera óptima para lograr una UX satisfactoria y efectiva.

3.2 Modelo de evaluación ITPX-ROBOTS

Partiendo de los fundamentos teóricos, se propone un nuevo modelo de evaluación para los contextos HRI en entornos industriales, denominado Individual Task Performance Experience – Robots (ITPX-ROBOT). El nuevo modelo ITPX-ROBOTS se compone de tres fases, de acuerdo con el enfoque holístico que deben tener las evaluaciones de la UX y de acuerdo con el modelo de evaluación ITPX (Mazmela, 2020). Las tres fases son (Figura 32):

- i) La evaluación experta previa, basada en heurísticos.
- ii) La evaluación durante la interacción.
- iii) La evaluación posterior a la interacción.

A continuación, se van a describir cada una de las fases (Figura 32).

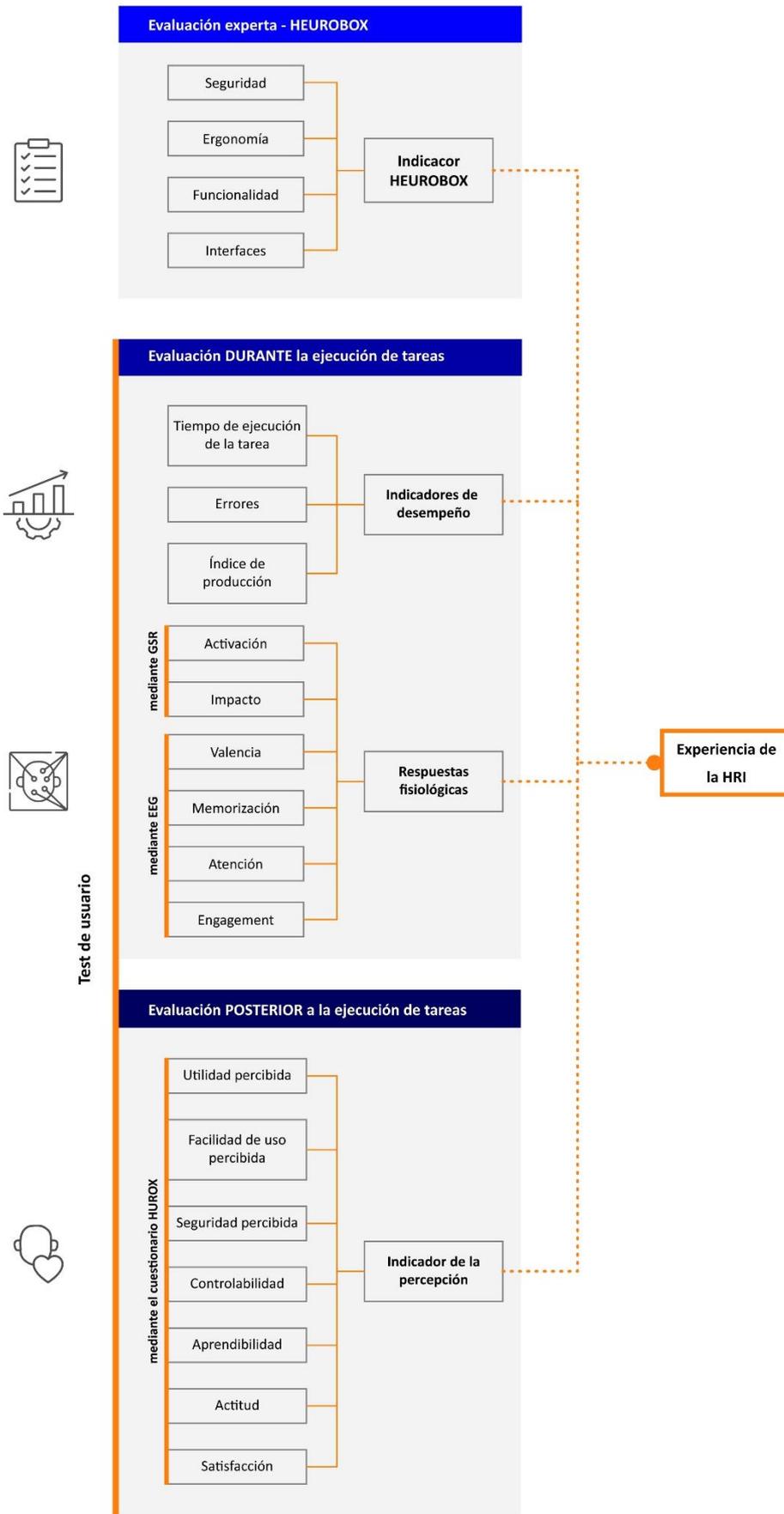


Figura 32: Modelo ITPX-Robots.

3.2.1 Evaluación experta

La evaluación experta planteada en el marco del nuevo modelo tiene como objetivo determinar si existe una compatibilidad entre las características de la persona, las del robot y las del entorno en cuestión. Este proceso de evaluación implica el uso de una serie de heurísticas, es decir, una lista de verificación o *checklist* que debe completar un experto en la evaluación de interfaces. Para llevar a cabo esta tarea, se ha desarrollado la herramienta heurística conocida como HEUROBOX, que se describe en detalle más adelante en el Capítulo IV.

HEUROBOX, por su parte considera como constructos la seguridad, la ergonomía, la funcionalidad y las interfaces (Figura 33).

- i) En cuanto a **la seguridad**, esta debe ser evaluada de manera exhaustiva para minimizar cualquier riesgo asociado al sistema en cuestión. Dichos riesgos abarcan aspectos relacionados con el proceso de fabricación, así como posibles peligros de naturaleza mecánica. Además, es imperativo que el sistema carezca de características técnicas y comportamientos que puedan generar lesiones. Por lo tanto, la evaluación experta debe asegurarse de que se implementen mecanismos de seguridad adecuados, así como la configuración de trayectorias que eviten cualquier tipo de atrapamiento del cuerpo humano, cables en movimiento o partes expuestas.
- ii) En lo que respecta a **la ergonomía**, esta debe ser evaluada abordando tanto las demandas físicas como las cognitivas sobre el operador. Las demandas físicas, como la postura, el alcance y la fuerza, deben asegurar que el operador realice las tareas de manera cómoda y sin esfuerzo ni fatiga. Por otro lado, la ergonomía cognitiva se enfoca en evaluar las demandas mentales impuestas al operador, como la atención, la toma de decisiones, la memoria y la capacidad de adaptarse a condiciones cambiantes.
- iii) En el ámbito de **la funcionalidad**, se evalúan cinco subcategorías: sistema, información, tarea, manejo de errores y asistencia. El sistema debe ser evaluado en términos de la predictibilidad y consistencia de la interacción, la flexibilidad para permitir al operador personalizar y controlar la interacción, y la capacidad de respaldo al operador. La subcategoría de información se enfoca en evaluar cómo el sistema proporciona información clara, precisa y relevante al operador, así como en su facilidad de acceso. La subcategoría de tarea evalúa si el sistema ayuda al operador a completar la tarea de manera eficiente y efectiva, eliminando pasos innecesarios e interferencias. El manejo de errores se centra en la capacidad del sistema para identificar y comunicar errores al operador, así como en su capacidad para ayudar al operador a recuperarse de ellos. Por último, la asistencia evalúa la capacidad del sistema para proporcionar información y retroalimentación clara y útil, el uso de reglas estándar para el soporte en la toma de decisiones y la provisión de ayuda y documentación contextual.
- iv) La evaluación experta de **las interfaces** se compone de las características generales necesarias y de cada tipo de interfaz (visual, de voz, gestual y háptica) que puede componer la HRI. Esta evaluación garantiza que la UI sea fácil e intuitiva para que la persona usuaria la comprenda y utilice, y que ayude al usuario a mantener una adecuada conciencia del estado del sistema, seguir la ejecución de tareas y tomar decisiones informadas sobre su interacción con el robot. En el caso de la interfaz visual, se evalúa el uso de gráficos e iconos simples, la minimización de ventanas múltiples, el uso de lenguaje y conceptos familiares, la usabilidad, accesibilidad y estética del diseño de la interfaz, la presentación adecuada de la información del sensor, la capacidad de manipular y almacenar la información mostrada, el uso de un lenguaje de interacción eficiente, y la consideración del color, contraste y visibilidad en el diseño de la interfaz. En el caso de la interfaz de voz, se evalúa el número de pasos

requeridos en el diálogo usuario-sistema, la precisión en la minimización de errores de entrada, el habla natural y similar al humano, el uso de un lenguaje de interacción eficiente y el diálogo estructurado entre la persona usuaria y el sistema. En el caso de la interfaz gestual, se evalúa que los gestos sean intuitivos y naturales para la persona usuaria y que puedan utilizarse para controlar los movimientos o acciones del robot. En cuanto a la interfaz háptica, se evalúa que el sistema pueda proporcionar retroalimentación táctil suficiente al usuario mediante el uso de objetos hápticos o táctiles, y que la persona usuaria pueda percibir e interpretar esta retroalimentación correctamente. Estas heurísticas se centran en la interpretabilidad y percepción de direcciones básicas, así como en los atributos táctiles de los objetos utilizados, como textura, fuerza, vibración, duración y aceleración.



Figura 33: Constructos que definen la evaluación experta.

3.2.2 Evaluación DURANTE la ejecución de tareas

La fase "durante" implica la realización de un test de usabilidad en el cual se recopilarán indicadores de rendimiento y valores fisiológicos. Para llevar a cabo este test de usabilidad, es necesario seguir las pautas y directrices establecidas en el protocolo desarrollado por Apraiz, Lasa, Montagna, et al. (2023), que proporcionan un marco de trabajo sólido y específico para llevar a cabo el test de manera estandarizada y objetiva (Figura 34).

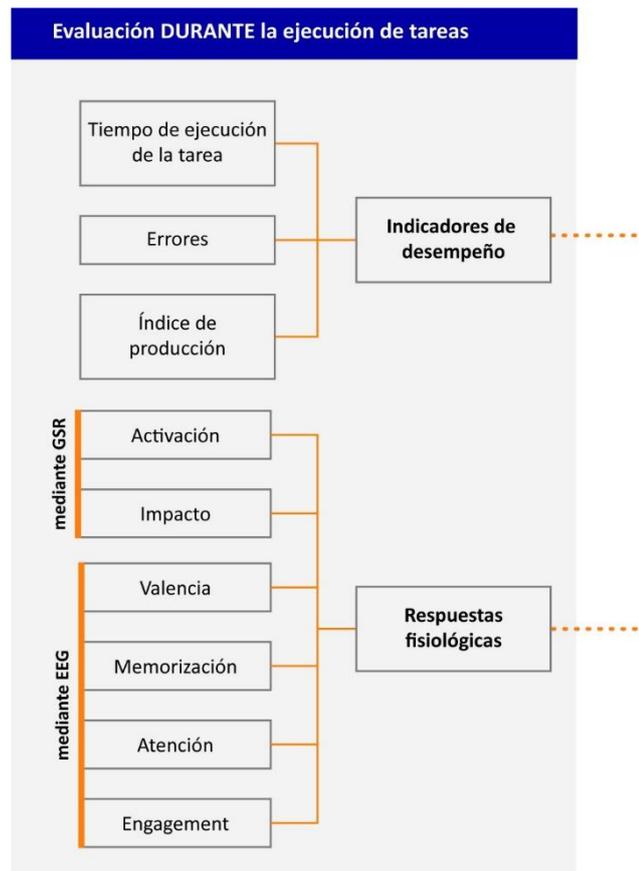


Figura 34: Evaluación DURANTE la ejecución de las tareas en el modelo ITPX-ROBOTS.

3.2.2.1 Indicadores de desempeño

Los indicadores de desempeño que se recogen en el modelo ITPX-ROBOTS son:

- i) **Tiempo de ejecución de la tarea:** Se refiere al período de tiempo necesario para que un usuario complete una tarea específica en la interacción con el sistema.
- ii) **Errores:** Las acciones incorrectas cometidas por las personas usuarias durante la realización de una tarea. Estos errores pueden incluir acciones no deseadas, selecciones incorrectas o pasos mal ejecutados. La métrica de errores se utiliza para evaluar la precisión y la capacidad del sistema para guiar a las personas usuarias a través de las tareas sin cometer errores. Menos errores suelen indicar una mejor usabilidad.
- iii) **Índice de producción:** El índice de producción es una métrica que evalúa la eficiencia de la producción o la finalización de tareas en relación con un estándar o referencia establecidos. Puede expresarse como un número o un porcentaje. Un índice de producción más alto indica que la persona usuaria está completando tareas a un ritmo más rápido y eficiente en comparación con el estándar establecido, lo que generalmente es un indicador positivo de la usabilidad y eficacia del sistema.

3.2.2.2 Indicadores fisiológicos

Además, el modelo ITPX-ROBOTS incorpora la evaluación de medidas fisiológicas. Concretamente, en el marco de este modelo, se emplean el dispositivo Ring y el EEG de Bitbrain. Estas herramientas fisiológicas de Bitbrain nos permiten obtener información sobre las siguientes biométricas:

- i) **Activación (Ring):** Este parámetro representa el nivel basal de activación fisiológica inducido por un estímulo o situación específica. Los valores inferiores a 0 sugieren un

estado de relajación o calma, mientras que los valores por encima de 0 indican un estado de excitación. Es importante destacar que la activación emocional puede estar relacionada tanto con emociones positivas como negativas (Bitbrain, 2020).

- ii) **Impacto (Ring):** El impacto emocional mide la cantidad y la intensidad de los cambios momentáneos en el estado emocional que resultan de una estimulación, un evento externo o la realización de una tarea. Se refiere a situaciones que llaman especialmente la atención, generan excitación o estrés (Bitbrain, 2020).
- iii) **Valencia (EEG):** Grado de atracción experimentado durante la presencia de un estímulo o una situación. Varía desde un extremo “positivo/agradable” hasta un extremo “negativo/desagradable” (Bitbrain, 2020).
- iv) **Memorización (EEG):** Este indicador mide la intensidad de los procesos cognitivos relacionados con la formación de futuros recuerdos durante la presentación de un estímulo. Captura el grado de almacenamiento, codificación y retención en la memoria (Bitbrain, 2020).
- v) **Atención (EEG):** Este indicador mide la focalización (hablando siempre en términos de procesos neuronales) de un participante cuando se le presenta un estímulo. Es decir, la utilización de los recursos cognitivos para la realización de una tarea o la visualización de un estímulo (Bitbrain, 2020).
- vi) **Engagement (EEG):** Es un indicador del grado de implicación o conexión entre el participante y el estímulo o tarea. No es simple atención ya que el participante puede estar totalmente atento y que la información que se presenta no sea de su interés y por lo tanto no exista dicha conexión o involucración (Bitbrain, 2020).

3.2.3 Evaluación POSTERIOR a la ejecución de las tareas

En la **fase posterior** a la ejecución de las tareas, el modelo ITPX-ROBOTS propone recoger la percepción de las personas usuarias. Para ello, el modelo propone utilizar el cuestionario HUROY, que se define en detalle más adelante en el Capítulo V. El cuestionario recoge los siguientes constructos (Figura 35):

- i) **Utilidad percibida:** El grado en que una persona cree que el uso de un sistema en particular mejoraría su desempeño laboral (Davis, 1989)
- ii) **Facilidad de uso percibida:** El grado en que una persona cree que el uso de un sistema en particular sería libre de esfuerzo y fácil de aprender (Davis, 1989).
- iii) **Seguridad percibida:** El grado en que una persona percibe que un sistema en particular puede adquirir nuevos conocimientos o habilidades (Akalin et al., 2022).
- iv) **Controlabilidad:** El grado en que una persona se siente en control de una tecnología y sus acciones, y la capacidad para modificar su comportamiento según sus necesidades o preferencias (Sarker & Wells, 2003).
- v) **Aprendibilidad:** El grado en que las personas usuarias pueden realizar con éxito una tarea cuando se encuentran con una interfaz por primera vez, así como el grado en que pueden volverse competentes en esa tarea con repeticiones crecientes de uso (Joyce, 2019).
- vi) **Actitud:** La evaluación general positiva o negativa de un usuario hacia una tecnología, componentes afectivos y cognitivos. Es un predictor de la intención de uso de la persona usuaria (Davis, 1989; Fishbein & Ajzen, 1977).
- vii) **Satisfacción:** El grado en que una persona está satisfecha con el uso de una tecnología, como resultado de la utilidad, facilidad de uso y confianza en su conjunto (K. C. Lee & Chung, 2009).

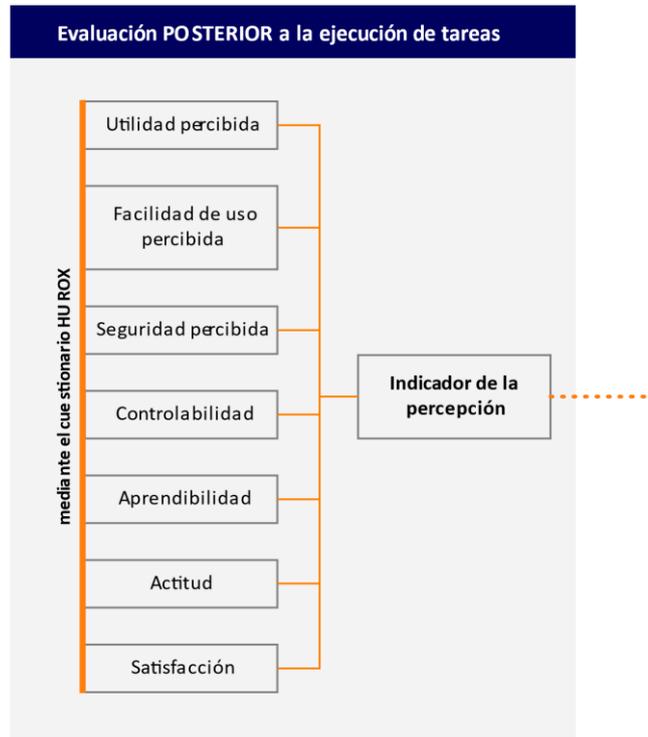


Figura 35: Evaluación POSTERIOR a la ejecución de las tareas en el modelo ITPX-ROBOTS.

Principales conclusiones del capítulo

- Se introduce un **nuevo marco conceptual llamado *Human Robot Interaction eXperience (HRIX)***, basado en el modelo CUE desarrollado por Thüning & Mahlke (2007). El marco sostiene que la interacción entre una persona y un robot se ve influenciada por tres conjuntos: (i) **las características de la persona**, (ii) **las características del robot**, y (iii) **las características del contexto**. Estos factores conjuntos determinan el grado de aceptación y la calidad de la interacción, y se denomina el **HRI FIT**. Este ajuste entre las características tiene un impacto en los componentes de la UX, que en este marco se definen como la **utilidad, la facilidad de uso, el confort, la aprendibilidad y la controlabilidad**. Todo ello, tiene unas consecuencias en los factores humanos como la actitud e intención de uso, la confianza, la seguridad percibida y la satisfacción.
- Se presenta la **taxonomía *Human-Robot-Context (HUBOXT)***, que señala que la aceptación de la HRI por parte de las personas trabajadoras se ve influenciada por una compleja interacción de diferentes factores. En primer lugar, las **características de la persona** que ejercen influencia en la aceptación de la HRI son (i) las características sociodemográficas, (ii) la motivación, (iii) la autoevaluación, (iv) las experiencias previas y el conocimiento, (v) las normas sociales subjetivas, (vi) la posición en el equipo, (vii) la personalidad y los valores morales y (viii) la intención de comportamiento. Por otro lado, **las características del robot** que ejercen influencia en la aceptación de la HRI son: (i) los atributos de seguridad, (ii) la tarea, (iii) la programación, (iv) el nivel de procesamiento y la memoria, (v) la precisión y repetibilidad, (vi) la carga útil y el alcance, (vii) la velocidad y la aceleración, (viii) la autonomía y el control, y, (ix) la flexibilidad y adaptabilidad. Por último, **las características del contexto** que impactan en la aceptación de la HRI se desglosan en dos niveles: las características del micro contexto y las características del macro contexto. Las características del micro contexto son: (i) el diseño físico del espacio de trabajo, (ii) la disposición de la tecnología, (iii) la carga de trabajo, y (iv) el proceso de implementación. Las características del macro contexto son: (i) la cultura organizacional, (ii) el liderazgo y la gestión, (iii) la relación con ejecutivos, y (iv) la relación con los compañeros y las compañeras.
- Se desarrolla el **modelo de evaluación *Individual Task Performance eXperience-ROBOTS (ITPX-ROBOTS)***, que adopta un enfoque integral para evaluar la UX en HRI en entornos industriales. El modelo se estructura en tres etapas fundamentales: (i) la evaluación previa, (ii) la evaluación durante la interacción y (iii) la evaluación posterior a la interacción.

Capítulo IV

Desarrollo y validación de la herramienta de evaluación heurística HEUROBOX

4 DESARROLLO Y VALIDACIÓN DE LA HERRAMIENTA DE EVALUACIÓN HEURÍSTICA HEUROBOX

La evaluación heurística es un método muy utilizado basado en las críticas de las personas evaluadoras para la rápida identificación de problemas de diseño (Molich & Nielsen, 1990) debido a su sencillez, bajo coste y amplia aplicabilidad. El uso iterativo de la heurística es fundamental para mejorar el rendimiento y la eficiencia de los espacios de trabajo HRI (Clarkson & Arkin, 2007), así como el bienestar de los operarios. Mediante el uso de heurísticas, es posible identificar la manera óptima en la que los robots pueden acercarse y colaborar con las personas trabajadoras, permitiendo que ambos colaboren por un mismo objetivo común.

Nielsen & Molich (1990) introdujeron la evaluación heurística como un método informal de usabilidad en el que múltiples evaluadores valoran el diseño de una interfaz y proporcionan comentarios basados en un conjunto de heurísticas desarrolladas para identificar posibles problemas de la interfaz. Como enumeran Quiñones & Rusu (2017), la evaluación heurística se puede aplicar en diferentes etapas del proceso de diseño y tiene varias ventajas, que incluyen: (i) bajo coste, (ii) no requiere una planificación exhaustiva, (iii) amplia aplicabilidad, especialmente en las primeras etapas del proceso de diseño, y (iv) capacidad para identificar problemas de usabilidad sin necesidad de personas usuarias. Sin embargo, también señalaron las siguientes desventajas: (i) los evaluadores deben tener experiencia y conocimientos adecuados para evaluar el producto; (ii) los evaluadores pueden no entender las tareas que realiza el producto, por lo que puede ser difícil identificar los problemas de usabilidad; y (iii) los problemas de usabilidad se identifican sin dar directamente una idea de cómo resolverlo.

Los conjuntos tradicionales de heurísticos son los 11 heurísticos de Nielsen, los 7 principios de usabilidad de (D. A. Norman, 2004), y los 8 criterios ergonómicos principales. A partir de ahí, los investigadores han desarrollado nuevas heurísticas para distintos contextos. En el contexto industrial, Mazmela (2020) y Aranburu (2020) han desarrollado listas de heurísticas específicas para aplicaciones industriales en interfaces hombre-máquina (HMI).

En cuanto a heurísticos o guías específicas para los contextos HRI, Qbilat et al. (2021) propusieron unas guías de accesibilidad. Además, en el estudio de Gualtieri, Rauch, et al. (2020) clasificaron los requisitos previos y las directrices de diseño para ayudar a los diseñadores a crear lugares de trabajo de ensamblaje colaborativo seguros, centrados en las personas y eficientes. En este sentido, Murphy & Tadokoro (2019) propusieron 32 directrices para establecer proactivamente buenas interfaces persona-robot. Frijns & Schmidbauer (2021) propusieron un conjunto de directrices para el diseño de la UI para la colaboración.

En el mismo contexto, Benitez Sandoval et al. (2018), propusieron principios de diseño para definir las HRI que incluían prácticas de programación, características de los robots como productos, comportamientos interactivos e incluso valores morales de la persona usuaria. Además, Bauer et al. (2009) proporcionaron reglas heurísticas para que los robots pidan direcciones en HRI, lo que resulta en una mejor interacción al reducir las ambigüedades, mejorar la intuición y permitir que el robot construya una representación interna de la ruta. Mediante el desarrollo de un nuevo conjunto de heurísticas para entornos industriales de HRI, las sugerencias en las primeras fases para mejorar los robots pueden suponer un ahorro de costes y tiempo (Weiss et al., 2010).

4.1 Dimensiones a considerar para un nuevo listado de heurísticos

Si bien existen varias listas de heurísticas disponibles para evaluar la robótica o las HMI desde diferentes perspectivas, hay una laguna en la evaluación heurística específica de los entornos HRI industriales desde un enfoque holístico centrado en las personas y en la UX. A diferencia de las experiencias con interfaces concretas, los entornos robóticos colaborativos son multimodales, lo que significa que la interacción se produce a través de múltiples canales. Como tales, todos estos canales deben proporcionar experiencias fluidas y satisfactorias. Además, la seguridad es de suma importancia en estos entornos de riesgo. En este contexto, el comportamiento del robot influye en la seguridad, la comodidad y la aceptación de la persona en el sistema robótico. A continuación, se describen las dimensiones a considerar para un nuevo listado de heurísticas para evaluar la HRI en entornos industriales.

- i) **Seguridad, confianza y seguridad percibida.** Garantizar la seguridad es el reto clave en el diseño y la implementación de HRI (Bänziger et al., 2020; Haddadin et al., 2011; Villani et al., 2018). El objetivo intrínseco de los sistemas colaborativos es facilitar la interacción directa entre personas y robots. Para que un equipo persona-robot alcance su objetivo, las personas deben confiar en que el robot protegerá los intereses y el bienestar de todos los demás individuos del equipo (Hancock et al., 2011). La confianza es especialmente importante, ya que afecta directamente a la disposición del operario a aceptar la información de los robots, seguir sus sugerencias y beneficiarse de las ventajas que ofrecen los sistemas robóticos. En relación con la confianza, la seguridad percibida es un factor importante a tener en cuenta.
- ii) **Ergonomía física.** Los nuevos entornos robóticos deben ser capaces de prevenir las molestias de las personas trabajadoras, así como el agravamiento de enfermedades relacionadas con las condiciones posturales o físicas en los espacios de trabajo, ya que los trastornos musculoesqueléticos constituyen la mayor categoría de enfermedades relacionadas con el trabajo en muchos países industriales (Punnett & Wegman, 2004).
- iii) **Ergonomía cognitiva y la hedonomía.** Los aspectos experienciales influyen en la aceptación del sistema, es decir, la usabilidad, la seguridad y el riesgo percibidos, las emociones y la percepción de la apariencia. Hay que tener en cuenta factores como la carga mental, el estrés físico y mental, la facilidad de aprendizaje, la utilidad, la eficiencia y el uso intuitivo. La seguridad y el riesgo percibido, aunque las emociones son de gran importancia en este contexto, ya que se consideran primordiales para una satisfacción óptima (Meissner et al., 2020; Villani et al., 2018). En cuanto a las emociones, hay que tener en cuenta la fiabilidad, la estimulación y la confianza (Meissner et al., 2020). Además, los aspectos hedonómicos como el atractivo y la estética visual del robot son cruciales. Estos factores influyen en el rendimiento y la satisfacción individual.
- iv) **Inclusividad.** La cuestión de la inclusividad se plantea en el ámbito de la robótica, donde no todas las personas usuarias pueden acceder a los robots debido a sus diversas características, como limitaciones visuales, auditivas, motoras o cognitivas. El descuido de estos aspectos durante la fase de diseño, implementación o interacción crea barreras de accesibilidad, impidiendo a las personas usuarias con discapacidades utilizar el sistema robótico (Qbilat et al., 2021).
- v) **Multimodalidad.** En los robots colaborativos, el uso de múltiples interfaces es habitual, lo que hace necesaria una nueva lista de heurísticas que tengan en cuenta este aspecto. Estas interfaces pueden permitir diversos modos de comunicación, como la comunicación gráfica, por voz o por gestos, y pueden variar en función del tipo de dispositivo utilizado. Prati et al. (2021) clasificaron las interfaces persona-robot en cuatro categorías principales: Pantallas visuales (por ejemplo, interfaces gráficas de usuario, interfaces de realidad aumentada), gestos (por ejemplo, movimientos faciales y de las manos), habla y lenguaje natural (por ejemplo, respuestas auditivas y basadas en

texto) e interacciones físicas y hápticas. Sin embargo, en la actualidad se carece de directrices establecidas para el control multimodal, lo que requiere una mayor investigación.

- vi) **Tipo de robots y funcionalidades.** En un contexto industrial, los robots pueden desarrollar diferentes tipos de acciones y tareas. Tal y como expusieron Mazmela Etxabe (2020) y Aranburu Zabalo (2020) en sus listas de heurísticas para HMIs industriales, las diferenciaron en función de las funcionalidades desempeñadas por el *software* específico. En este sentido, es posible extrapolarlas al contexto de la robótica industrial. Así, sería conveniente que la lista estuviera compuesta por diferentes tipologías de robots y funcionalidades.

Por lo tanto, se ha identificado la necesidad de crear una nueva lista de heurísticos para evaluar la HRI desde un enfoque de UX, aceptación de la tecnología y centrado en el ser humano, que tenga en cuenta la seguridad, la confianza y la seguridad percibida, la ergonomía física, la ergonomía cognitiva y las emociones, la inclusividad, la multimodalidad y los tipos de robots y funcionalidades.

4.2 Metodología

La metodología se basa en establecer y orientar características específicas de la aplicación, identificar heurísticas existentes para su reutilización -cuando sea posible-, especificarlas según la plantilla y, finalmente, su validación (Quiñones & Rusu, 2017). Cuando no se dispone de heurísticas y es necesario desarrollar nuevas heurísticas, las recomendaciones y guías de diseño para la aplicación específica pueden ser de valioso interés. La Figura 36 muestra brevemente la metodología seguida para establecer y validar la lista de heurísticas propuestas para la HRI.

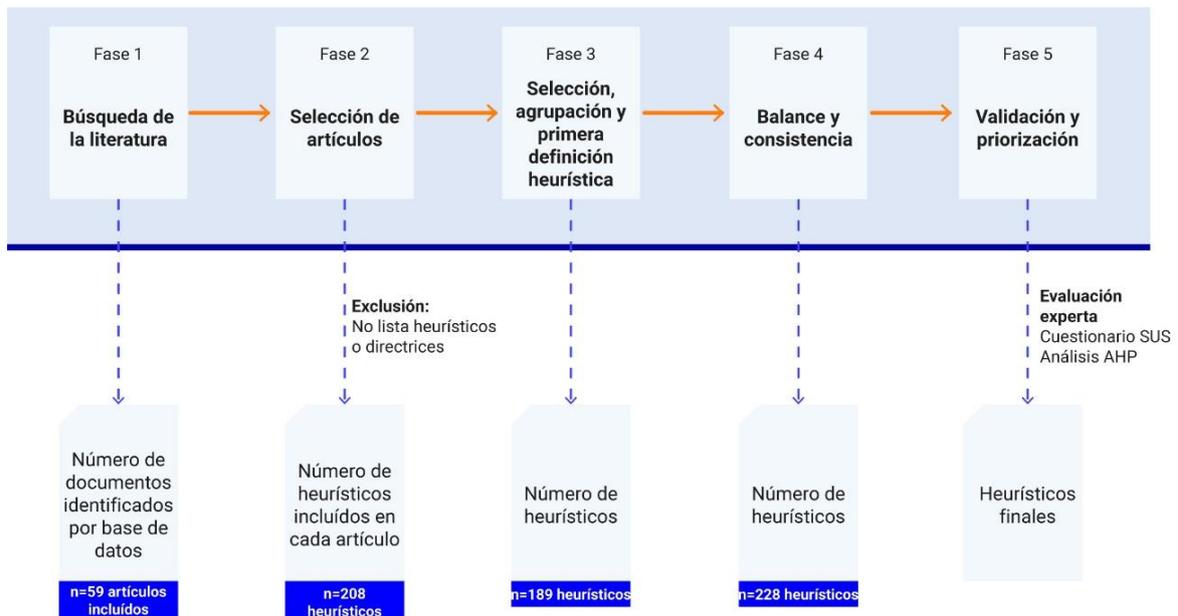


Figura 36: Metodología para el desarrollo de un nuevo método de evaluación heurística de la HRI.

4.2.1 Fase 1: Búsqueda de la literatura

La primera fase consiste en recopilar heurísticas o directrices de HRI existentes en diferentes bases de datos. La Tabla 28, muestra las bases de datos seleccionadas para realizar esta revisión.

Tabla 28: Bases de datos seleccionadas para la búsqueda de listados de heurísticos o guías en la literatura.

Base de datos	Tipo de base de datos	Descripción
Scopus	Base de datos de citas	Servicio de indexación de citas científicas para la búsqueda de citas de artículos de revistas revisadas por expertos. Se utiliza principalmente en los cálculos bibliométricos, de Elsevier.
ACM Digital Library	Base de datos de investigación	Base de datos científica en materias relacionadas con la informática y ciencias computacionales.
Engineering Village	Base de datos de investigación	Base de datos especializada en ingeniería.
Science Direct	Base de datos de la editorial	Ofrece sus propias revistas científicas a texto completo. Cubre la mayoría de las disciplinas, pero con un enfoque en la ciencia, la tecnología y ciencias sociales, de Elsevier y otras editoriales relacionadas.

Para ello, se llevó a cabo una revisión sistemática de la bibliografía. Se creó la siguiente ecuación de búsqueda. No obstante, cada base de datos requiere la adaptación de la ecuación en función de las diferencias en sus funcionalidades, como se detalla en la Tabla 29.

("heuristic evaluation" or "design guidelines") and ("human-robot interaction" or "hri" or "human-robot collaboration" or "hrc")

Tabla 29: Ecuación de búsqueda por cada una de las bases de datos.

Base de datos	Fecha de búsqueda	Ecuación de búsqueda	N.º de resultados
Scopus	03/06/2022	((heuristic AND evaluation) AND (((human-robot AND interaction) OR hri) OR ((human-robot AND collaboration) OR hrc)))	27
ACM Digital Library	06/06/2022	[[Abstract: "heuristic evaluation"] OR [Abstract: or] OR [Abstract: "design guidelines"] OR [Abstract: and] OR [Abstract: "human-robot interaction"] OR [Abstract: or "hri" or] OR [Abstract: "human-robot collaboration"] OR [Abstract: or "hrc"]] AND [[Title: "heuristic evaluation"] OR [Title: or] OR [Title: "design guidelines"] OR [Title: and] OR [Title: "human-robot interaction"] OR [Title: or "hri" or] OR [Title: "human-robot collaboration"] OR [Title: or "hrc"]]	2
Engineering Village	06/06/2022	((("heuristic evaluation" OR "design guidelines") AND ("human-robot interaction" OR "HRI" OR "human-robot collaboration" OR "HRC")) WN AB) + (2021 OR 2020 OR 2019 OR 2018 OR 2017 OR 2016 OR 2015) WN YR	20
Science direct	06/06/2022	Title, abstract, keywords: ("heuristic evaluation" OR "DESIGN GUIDELINES") AND ("HUMAN-ROBOT INTERACTION" OR "HRI" OR "HUMAN-ROBOT COLLABORATION" OR "HRC")	10
RESULTADOS			59

4.2.2 Fase 2: Selección de artículos

Tras identificar los trabajos anteriores, se realizó una revisión de la bibliografía (un total de 59 trabajos). Se recopilaron todas las heurísticas y directrices de diseño pertinentes en el contexto de la HRI, que inicialmente ascendieron a un total de 208 heurísticas tentativas, como puede verse en la Tabla 30.

Tabla 30: Identificación inicial de los heurísticos y guías.

Autor/es	Año	Título del artículo	No. de heurísticos o guías
Drury et al.	2004	Design guidelines for improved human-robot interaction	4
Clarkson and Arkin	2007	Applying heuristic evaluation to human-robot interaction systems	7
Weiss et al.	2010	A methodological adaptation for heuristic evaluation of HRI	7
Tsui et al.	2010	Developing heuristics for assistive robotics	29
Andonovski et al.	2010	Towards the development of a haptics guideline in human-robot systems	22
Keebler et al.	2012	Applying Team Heuristics to Future Human-Robot Systems	5
Adamides et al.	2014	Usability Guidelines for the Design of Robot Teleoperation: A Taxonomy	8
Manning et al.	2015	Heuristic Evaluation of Swarm Metrics' Effectiveness	7
Coronado et al.	2017	Gesture-based robot control: Design challenges and evaluation with humans	6
Gualtieri, Monizza, et al.	2020	From Design for Assembly to Design for Collaborative Assembly - Product Design Principles for Enhancing Safety, Ergonomics and Efficiency in Human-Robot Collaboration	16
Gualtieri, Rauch, et al.	2020	Safety, Ergonomics and Efficiency in Human-Robot Collaborative Assembly: Design Guidelines and Requirements	16
Adamides, 2020	2020	Heuristic Evaluation of the User Interface for a Semi-Autonomous Agricultural Robot Sprayer	8
Qbilat et al.	2021	A Proposal of Accessibility Guidelines for Human-Robot Interaction	18
Frijns and Schmidbauer,	2021	Design Guidelines for Collaborative Industrial Robot User Interfaces	24
Gualtieri et al.	2022	Development and validation of guidelines for safety in human-robot collaborative assembly systems	31
TOTAL			208

4.2.3 Fase 3: Selección, agrupación y primera definición heurística

Debido a la complejidad de la interacción, en la bibliografía se tratan múltiples temas en cuanto a evaluación heurística y directrices de diseño, es decir, funcionalidades, interfaces, ergonomía, etc. Para ello, agrupamos y clasificamos todas las frases tentativas recopiladas en el paso anterior. Estos grupos se definen estratégicamente por su naturaleza e implicaciones en la interacción para ayudar y guiar fácilmente a los evaluadores que puedan utilizar esta herramienta (Tabla 31).

Tabla 31: Definición de categorías para la clasificación de directrices de diseño y heurísticas.

Categorías	Subcategorías	Definición
Seguridad	General	Evalúa aspectos genéricos de la seguridad del sistema, como el riesgo de lesiones o daños a las personas usuarias, el cumplimiento de las normas de seguridad y la capacidad del sistema para responder a emergencias.
	Planificación del movimiento	Evalúa la capacidad del sistema para planificar y ejecutar movimientos seguros en el espacio de trabajo, como evitar colisiones con otros objetos o personas, mantener una distancia de seguridad con los peligros y optimizar los movimientos en aras de la eficacia y la seguridad.
	Sistemas del robot	Analiza las características del sistema robótico y los espacios de trabajo, incluido el tipo de robot utilizado, sus capacidades y limitaciones, la disposición física del espacio de trabajo y la presencia de otros peligros u obstáculos.
	Medidas organizacionales	Analiza las medidas organizativas existentes para garantizar unas condiciones seguras, como los programas de formación y certificación de las personas usuarias, los protocolos y procedimientos de seguridad y la presencia de personal o equipos de seguridad.
Ergonomía	Ergonomía física	Se refiere a las características anatómicas, antropométricas, fisiológicas y biomecánicas humanas en relación con la actividad física (Asociación Internacional de Ergonomía, 2019), como la postura, el alcance, la fuerza de agarre y la amplitud de movimiento.
	Ergonomía cognitiva	Cubre la adecuación del uso a las capacidades cognitivas de las personas usuarias, incluidas la percepción humana, el procesamiento mental y la memoria, para garantizar que el sistema sea fácil de aprender, utilizar y recordar.
Funcionalidad	Sistema	Cubre la facilidad de uso del sistema en términos de funciones disponibles, como la gama de tareas que el sistema puede realizar, la facilidad de acceso y uso de estas funciones y el nivel de personalización disponible para las personas usuarias.
	Información	Define la evaluación de cómo se presenta la información del sistema a las personas usuarias, como la claridad, relevancia y accesibilidad de la información, y el grado en que ayuda a las personas usuarias a realizar sus tareas.
	Tarea	Se basa en la eficacia de la tarea y en lo bien que se ejecuta, como la velocidad y precisión de la ejecución de la tarea, el nivel de compromiso y satisfacción de la persona usuaria y la capacidad de recuperación ante errores o interrupciones.
	Gestión de errores	Metodologías para descubrir, capturar y recuperarse de errores en el sistema, como mecanismos de detección y corrección de errores, retroalimentación y notificación al usuario, y estrategias automatizadas de recuperación o compensación.
	Asistencia	Define la documentación o las metodologías de ayuda del sistema, como manuales de usuario, sistemas de ayuda en línea, orientación contextual y tutoriales o demostraciones en el sistema.
Interfaces	General	Implica la evaluación general de la interfaz, así como la coherencia y claridad de los elementos e interacciones de la interfaz.
	Visual	Analiza la interfaz basándose en la información mostrada en los visores o pantallas, como la claridad y legibilidad del texto y los gráficos, el uso del color y el contraste, y la jerarquía visual y organización general de la información.
	Gestual	Tiene en cuenta las entradas de gestos en el sistema, como la facilidad e intuitividad de las interacciones basadas en gestos, la coherencia y fiabilidad de estas interacciones en diferentes personas usuarias y contextos, y la capacidad de recuperarse de los errores de reconocimiento de gestos.

Háptica	Evalúe el rendimiento del sistema a través de interacciones táctiles, como la sensibilidad y precisión de las entradas táctiles, la retroalimentación proporcionada a través de vibraciones u otras señales táctiles, y la capacidad de personalizar o ajustar la configuración de la retroalimentación háptica.
Voz	Evalúa la información presentada y transmitida mediante sonidos y voz, como la claridad e inteligibilidad de las indicaciones o instrucciones vocales.

En este paso, todas las directrices se redefinen también como heurísticas. Este paso requirió un enorme esfuerzo para captar perfectamente el significado de la frase de evaluación, ya que la complejidad de las interacciones y las múltiples referencias podían fundirse en una sola heurística. Para cada subgrupo, las heurísticas se agrupan en dos niveles (básico y avanzado), atendiendo a la complejidad de la evaluación y a lo específico que sea el aspecto evaluado.

4.2.4 Fase 4: Balance y consistencia

Tras la obtención de una lista preliminar de resultados heurísticos derivados del paso anterior, es necesario evaluar la coherencia de cada categoría y subcategoría y esforzarse por lograr un equilibrio entre ellas. Tras esta iteración inicial en el proceso de desarrollo heurístico, se llevó a cabo una revisión bibliográfica específica para abordar cualquier incoherencia identificada. En la Tabla 32 se presenta un resumen de la bibliografía utilizada para equilibrar la heurística.

Tabla 32: Literatura heurística específica existente para equilibrar la nueva lista.

Autor/es	Año	Título	Tema	N. de heurísticos propuestos
Bauer et al.	2009	Heuristic Rules for Human-Robot Interaction Based on Principles from Linguistics-Asking for Directions	Principles from Linguistic	10
Wibowo et al.	2017	Heuristic Evaluation and User Testing with ISO 9126 in Evaluating of Decision Support System for Recommendation of Outstanding Marketing Officer	Decision Support System	10
Maguire	2019	Development of a heuristic evaluation tool for voice user interface	Voice Interface	8
Andonovski et al.	2010	Towards the development of a haptics guideline in human-robot systems	Haptic Interface	11
TOTAL				39

4.2.5 Fase 5: Validación y priorización

Una vez establecidas todas las heurísticas como un borrador inicial, se procedió a someter la lista a un grupo de evaluadores especializados. Estos evaluadores fueron seleccionados previamente en base a su formación específica y a sus roles o funciones profesionales, siguiendo los criterios de inclusión establecidos. En concreto, se consultó a un total de 15 evaluadores externos, considerando su experiencia profesional como factor determinante. Las invitaciones fueron enviadas mediante correos electrónicos formales, que detallaban el tema de investigación y los objetivos del estudio. Posteriormente, se les proporcionó a las personas expertas la encuesta de expertos en formato PDF, la cual forma parte del conjunto de datos (Apraiz, Mulet Alberola, et al., 2023).

Se utilizó el cuestionario SUS (Brooke, 1996) para medir la usabilidad percibida de la herramienta HEUROBOX bajo la evaluación de las personas expertas. Cada experto reclutado participó inicialmente en un estudio de usabilidad a través del cuestionario SUS, con el objetivo de evaluar la idoneidad de la evaluación heurística para su aplicación en el ámbito de la HRI. Esta técnica es ampliamente conocida y utilizada en diversos campos, como el diseño de ingeniería (Gopsill et al., 2015), la ingeniería del *software* (Ya-feng et al., 2022) o los sistemas de servicios inteligentes (Chang et al., 2019). Además, el SUS se puede aplicar en encuestas a pequeña escala, cuyos elementos del cuestionario pueden ser fácilmente comprendidos por un número limitado de participantes con diversas disciplinas (Martins et al., 2015).

Además, la encuesta incluyó un cuestionario de valoración de importancia en forma de comparación por pares, siguiendo el Proceso Analítico Jerárquico (AHP, por sus siglas en inglés). Este método se utilizó para solicitar a las personas expertas que calificaran el grado de importancia de cada una de las categorías de la herramienta HEUROBOX: Seguridad, Ergonomía, Funcionalidad e Interfaces. Estos grados de importancia permiten a profesionales e investigadores priorizar qué categoría es más importante que las demás, lo que les permite tomar decisiones adecuadas en la evaluación y mejora de sistemas persona-robot en entornos industriales. Seguimos el procedimiento de AHP proporcionado por Saaty (2004), codificado en R por Cho (2019), para evaluar la importancia de cada una de las categorías de la herramienta HEUROBOX en una escala de comparación por pares de 9 puntos, como se muestra en la Tabla 33.

Tabla 33: Presentación del método de evaluación AHP.

Esta categoría se prefiere sobre la categoría de la derecha	Igual									Esta categoría se prefiere sobre la categoría de la izquierda								
	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	
Elemento A																	Elemento B	

Los resultados de la encuesta realizada a las personas expertas, así como los datos sin procesar de las respuestas proporcionadas por ellos, se encuentran disponibles en el conjunto de datos publicado (Apraiz, Mulet Alberola, et al., 2023).

4.3 Listado completo de heurísticos

Esta sección contiene la lista de heurísticas generadas en base a la búsqueda bibliográfica y su posterior selección y agrupación. Como se puede observar en la Tabla 34, HEUROBOX se compone de cuatro categorías (Seguridad, Ergonomía, Funcionalidad e Interfaces). Además, en cada categoría se diferencian el nivel básico y el nivel avanzado. El nivel básico tiene como objetivo cubrir los aspectos esenciales de una evaluación HRI, e incluye un total de 84 principios heurísticos. El nivel avanzado abarca aspectos de elementos o funciones específicas, y suma un total de 228 (incluyendo los del nivel básico) principios heurísticos. A continuación, se explicará el desarrollo de la lista por categoría, pero es importante destacar que, para proporcionar una experiencia fluida y satisfactoria de manera holística, es importante evaluar cada categoría.

Tabla 34: El conjunto de heurísticas que componen HEUROBOX.

Categoría	Subcategoría	N. total de heurísticas	N. total de heurísticas en la evaluación básica	N. total de heurísticas en la evaluación avanzada
Seguridad		30	7	23
	General	7	7	0
	Planificación del movimiento	10	0	10
	Sistemas del robot	9	0	9
	Medidas organizacionales	4	0	4
Ergonomía		49	13	36
	Ergonomía física	29	4	25
	Ergonomía cognitiva	20	9	11
Funcionalidad		78	36	42
	Sistema	18	6	12
	Información	28	13	15
	Tarea	12	7	5
	Gestión de errores	14	7	7
	Asistencia	6	3	3
Interfaces		71	28	43
	General	28	28	0
	Visual	22	0	22
	Voz	11	0	11
	Háptica	6	0	6
	Gestual	4	0	4
TOTAL		228	84	144

4.3.1 Seguridad

La categoría de Seguridad en la herramienta de evaluación HEUROBOX comprende dos niveles: un nivel básico, que incluye una sección genérica con 7 heurísticas, y un nivel avanzado, que se divide en tres subsecciones: Planificación del movimiento (10 heurísticas), Sistemas del robot (9 heurísticas) y Medidas organizativas (4 heurísticas). La evaluación avanzada en Seguridad incluye un total de 30 heurísticas, que tienen como objetivo proporcionar una evaluación más completa de la seguridad de la HRI en entornos de ensamblaje. Las heurísticas de Seguridad se han nutrido de las heurísticas y directrices propuestas por Bauer et al. (2009), Gaultieri et al. (2022), Gaultieri, Monizza, et al. (2020), Gaultieri, Rauch, et al. (2020), Tsui et al. (2010), y Weiss et al. (2010).

En base a esta evaluación, el sistema debe estar diseñado para minimizar los riesgos relacionados con la HRI en entornos de ensamblaje. Estos riesgos incluyen riesgos generales relacionados con la interacción de las partes humano-ensamblaje, riesgos mecánicos específicos relacionados con la interacción de las partes humano-ensamblaje y riesgos mecánicos específicos relacionados con la caída de las partes del sistema del robot. El sistema también debe carecer de características o comportamientos físicos que puedan causar lesiones, e incluir mecanismos de seguridad para garantizar la seguridad. Además, el sistema debe permitir establecer trayectorias para evitar el contacto o atrapamiento de partes del cuerpo humano, y debe incluir funciones para limitar velocidades, fuerzas y torsiones. El sistema también debe contar con mecanismos de absorción de energía y reducción de fuerzas de impacto, así como sensores para anticipar o detectar el contacto y características de

protección contra riesgos asociados con la pieza de trabajo. Por último, el sistema debe incluir características para prevenir el atrapamiento debido a cables en movimiento o partes expuestas, y debe contar con señalización y resaltado para alertar a las personas operarias sobre posibles riesgos.

4.3.2 Ergonomía

Las heurísticas ergonómicas se organizan en dos subcategorías principales: (i) ergonomía física y (ii) ergonomía cognitiva. Cada una de estas subcategorías se divide además en dos niveles de evaluación: (i) nivel básico y (ii) nivel avanzado. En general, las heurísticas en Ergonomía proporcionan una evaluación integral de la ergonomía de la HRI en entornos de ensamblaje, abarcando tanto las demandas físicas como las cognitivas de la persona operaria. Las heurísticas de Ergonomía se basan en las directrices e investigaciones realizadas por Adamides et al. (2014), Drury et al., (2004), Frijns & Schmidbauer (2021), Gualtieri, Monizza, et al. (2020), T. Lewis et al. (2014), y Weiss et al. (2010).

- Las heurísticas en la subcategoría de **ergonomía física**, que consta de 4 heurísticas en el nivel básico y 25 en el nivel avanzado, se centran en las demandas físicas a las que se enfrenta el operador durante la HRI, como postura, alcance y fuerza. Estas heurísticas tienen como objetivo garantizar que el operador pueda realizar las tareas de manera cómoda y sin esfuerzo ni fatiga. El nivel avanzado de evaluación en esta categoría incluye heurísticas adicionales que abordan cuestiones más complejas.
- Las heurísticas en la subcategoría de **ergonomía cognitiva**, que consta de 9 heurísticas en el nivel básico y 11 en el nivel avanzado, se centran en las demandas mentales a las que se enfrenta el operador durante la HRI, como la atención, la toma de decisiones y la memoria. Estas heurísticas tienen como objetivo garantizar que el operador pueda procesar la información de manera efectiva y tomar decisiones mientras interactúa con el robot. El nivel avanzado de evaluación en esta categoría incluye heurísticas adicionales que consideran demandas cognitivas más avanzadas, como la capacidad de adaptarse a condiciones cambiantes y aprender de la experiencia.

4.3.3 Funcionalidad

Las heurísticas de Funcionalidad se organizan en cinco subcategorías principales: (i) sistema, (ii) información, (iii) tarea, (iv) manejo de errores y (v) asistencia. Cada una de estas subcategorías se divide además en dos niveles de evaluación: (i) nivel básico y (ii) nivel avanzado. Las heurísticas de la subcategoría Funcionalidad se basan en las directrices e investigaciones realizadas por Adamides et al. (2014), Clarkson & Arkin (2007), Drury et al. (2004), Frijns & Schmidbauer (2021), Gualtieri, Monizza, et al. (2020), Keebler et al. (2012); T. Lewis et al. (2014), Qbilat et al. (2021), Tsui et al. (2010), Weiss et al. (2010), Wibowo et al. (2017), y Young et al. (2011).

- La subcategoría de **Sistema** - 6 heurísticas en el nivel básico y 13 heurísticas en el nivel avanzado - abarca una amplia gama de cuestiones, incluyendo la previsibilidad y consistencia de la interacción, la flexibilidad del sistema para permitir al operador personalizar y controlar la interacción, y la capacidad del sistema para apoyar al operador.
- Las heurísticas en la subcategoría de **Información** - 13 heurísticas en el nivel básico y 15 heurísticas en el nivel avanzado - se centran en la forma en que el sistema proporciona información clara, precisa y relevante al operador, y en que permite al operador acceder a la información de manera conveniente y eficiente.

- Las heurísticas en la subcategoría de **Tarea** - 7 heurísticas en el nivel básico y 5 heurísticas en el nivel avanzado - tienen como objetivo garantizar que el sistema ayude al operador a completar la tarea de manera eficiente y efectiva, sin pasos innecesarios o interferencias. Estas heurísticas se centran en simplificar diversos tipos de tareas, como el reconocimiento de objetos, la alimentación, la manipulación y el ensamblaje, y en minimizar el número de pasos necesarios para alcanzar los objetivos. Las heurísticas también consideran la capacidad del sistema para adaptarse a diferentes tipos de tareas y escenarios de aplicación, y para permitir la reutilización de trabajos anteriores.
- Las heurísticas en la subcategoría de **Gestión de errores** - 7 heurísticas en el nivel básico y 7 heurísticas en el nivel avanzado - se centran en la capacidad del sistema para identificar y comunicar errores al operador, así como en su capacidad para ayudar al operador a recuperarse de los errores. Esto incluye el uso de mecanismos visuales para indicar errores, proporcionar mensajes de error claros y comprensibles, y ofrecer opciones para la recuperación o reversión de acciones. Las heurísticas también consideran la capacidad del sistema para autoinspeccionarse en busca de daños u obstáculos, y para proporcionar información sobre el entorno de la tarea para ayudar al operador a comprender la causa de los errores.
- Las heurísticas en la subcategoría de **Asistencia** - 3 heurísticas en el nivel básico y 3 heurísticas en el nivel avanzado - tienen como objetivo evaluar la asistencia proporcionada por el sistema. Cubren una serie de cuestiones relacionadas con la asistencia, incluyendo la capacidad del sistema para proporcionar información y retroalimentación clara y útil, el uso de reglas estándar para el apoyo a la toma de decisiones, y la provisión de ayuda y documentación contextualmente relevantes. Además, las heurísticas abordan la importancia de apoyar y ayudar al operador a determinar el nivel más apropiado de autonomía robótica en cualquier momento dado.

4.3.4 Interfaces

Estas heurísticas se organizan en dos niveles: (i) nivel básico, compuesto por cuestiones generales, y (ii) nivel avanzado, compuesto por cuatro subcategorías (visual, voz, gestos y háptica). Al evaluar un sistema con el nivel avanzado, el evaluador experto debe seleccionar solo las subcategorías que estén involucradas en ese contexto particular del sistema. Las heurísticas sobre Interfaces se basan en las directrices e investigaciones realizadas por Adamides et al. (2014), Andonovski et al. (2010), Bauer et al. (2009), Clarkson & Arkin (2007), Coronado et al. (2017), Drury et al. (2004), Frijns & Schmidbauer (2021), Fulfagar et al. (2021), Keebler et al. (2012), Maguire (2019), Qbilat et al. (2021), y Weiss et al. (2010).

- El nivel de evaluación básico - 28 heurísticas - proporciona una evaluación general de la UI en la HRI sin considerar el tipo de interfaz. Estas heurísticas tienen como objetivo garantizar que la UI sea fácil e intuitiva de entender y utilizar, y que ayude al usuario a mantener una conciencia adecuada del estado del sistema, seguir la ejecución de tareas y tomar decisiones informadas sobre su interacción con el robot.
- Las heurísticas para interfaces visuales - 22 heurísticas - tienen como objetivo garantizar que la UI visual del sistema HRI sea fácil de usar y comprender, y que presente la información necesaria de manera clara y estructurada. Estas heurísticas abarcan una amplia gama de temas, como el uso de gráficos e iconos simples, la minimización de ventanas múltiples, el uso de un lenguaje y conceptos familiares, la usabilidad, accesibilidad y estética del diseño de la

interfaz, la presentación adecuada de la información del sensor, la capacidad de manipular y almacenar la información mostrada, el uso de un lenguaje de interacción eficiente, y la consideración del color, contraste y visibilidad en el diseño de la interfaz.

- Las heurísticas para interfaces de voz - 11 heurísticas - tienen como objetivo garantizar que la UI de voz (VUI) del sistema sea fácil de entender y recordar, eficiente de usar y que proporcione comentarios y soporte adecuados al operador. Estas heurísticas abarcan una serie de temas, como el número de pasos requeridos en el diálogo entre la persona y el sistema, la precisión en la minimización de errores de entrada, el habla natural y similar a la humana, el uso de un lenguaje de interacción eficiente y el diálogo estructurado entre la persona usuaria y el sistema.
- Las heurísticas para gestos - 4 heurísticas - tienen como objetivo garantizar que los gestos sean intuitivos y naturales para la persona usuaria, y que se puedan utilizar para controlar los movimientos o acciones del robot.
- Las heurísticas para la interfaz háptica - 6 heurísticas - tienen como objetivo garantizar que el sistema pueda proporcionar suficiente retroalimentación táctil al usuario a través del uso de objetos hápticos o táctiles, y que la persona usuaria pueda percibir e interpretar correctamente esta retroalimentación. Estas heurísticas se centran en la interpretación y percepción de las direcciones básicas, así como en los atributos táctiles de los objetos utilizados, como textura, fuerza, vibración, duración y aceleración.

4.4 Resultados de la validación y priorización de las personas expertas

Esta sección se centra en presentar los resultados de validación y priorización realizados por las personas expertas, siguiendo la metodología definida. Nguyen Ngoc et al. (2022) resaltaron una brecha crucial en la literatura existente en relación a las herramientas de diseño: la necesidad de evaluaciones empíricas de usabilidad. Esto motivó a abordar esta limitación mediante la realización de una evaluación de la herramienta HEUROBOX. Para lograr este objetivo, se siguieron las indicaciones de Nguyen Ngoc et al. (2022) para realizar una evaluación que involucró la obtención de opiniones de expertos utilizando el cuestionario SUS y la metodología AHP. El conjunto de datos presentado se encuentra almacenado en Mendeley Data (Apraiz, Mulet Alberola, et al., 2023).

Como se expuso previamente, un total de 15 expertos participaron en la evaluación de HEUROBOX (Tabla 35), incluyendo seis mujeres (40%), una persona no binaria (6,7%) y una persona que optó por no revelar su género (6,7%). De las personas participantes, dos eran profesionales (13,3%), mientras que el 86,7% restante eran académicos.

Tabla 35: Listado de expertos que participó en la validación de HEUROBOX.

N. de experto	Sexo	Conocimientos	Campos principales	Años trabajando
1	Hombre	Academia	Ingeniería biomédica, biomecánica, HRI y aprendizaje por refuerzo	3
2	Hombre	Academia	Informática, inteligencia artificial, aprendizaje automático e IoT	14
3	Hombre	Academia	Ingeniería industrial, Diseño para la fabricación, Diseño para el montaje	4
4	Mujer	Academia	Psicología, UX en RV	5
5	Mujer	Academia	Ingeniería biomédica, tecnologías digitales y	9

			diseño centrado en el usuario	
6	Hombre	Profesional	Sistemas robóticos, PLC y conocimientos	1
7	No binario	Academia	Ingeniería electrónica, control y automatización	3
8	Prefirió no contestar	Academia	Ingeniería electrónica, control	11
9	Hombre	Profesional	Sistemas robóticos, automatización, Máquina herramienta	15
10	Mujer	Academia	Experiencia del usuario, aceptación de la tecnología, HRI	4
11	Hombre	Academia	Mecatrónica, aeroespacial, física, economía circular	4
12	Mujer	Academia	Diseño inclusivo, Brecha digital	3
13	Mujer	Academia	Experiencia del usuario, aceptación tecnológica, interfaz persona-máquina	12
14	Hombre	Academia	Experiencia de usuario, Realidad virtual	4
15	Mujer	Academia	Diseño centrado en las personas	3

4.4.1 Evaluación de la usabilidad

Como indicaron Nguyen et al. (2022), para medir la usabilidad de HEUROBOX, se utilizó una versión simplificada del SUS (Brooke, 1996), un cuestionario de 10 ítems que mide la percepción de usabilidad aplicando opciones de respuesta Likert de 5 puntos (totalmente en desacuerdo a totalmente de acuerdo).

El SUS consta de 10 preguntas, donde las preguntas impares tienen significados positivos y las preguntas pares tienen significados negativos. Las preguntas positivas se puntúan restando un punto a la puntuación de la persona usuaria, mientras que las preguntas negativas se puntúan restando la puntuación de la persona usuaria a 2. Las puntuaciones totales se multiplican por 2,5 para obtener un rango de 0 a 100. La puntuación promedio del SUS para HEUROBOX es de 76 desde la perspectiva de las personas expertas. Según se indica en el rango de adjetivos de las puntuaciones del SUS por Bangor et al. (2009), las calificaciones de HEUROBOX se clasifican como "excelentes". Los resultados en la Figura 37 representan las calificaciones promedio otorgadas por las personas expertas para cada ítem del cuestionario SUS.

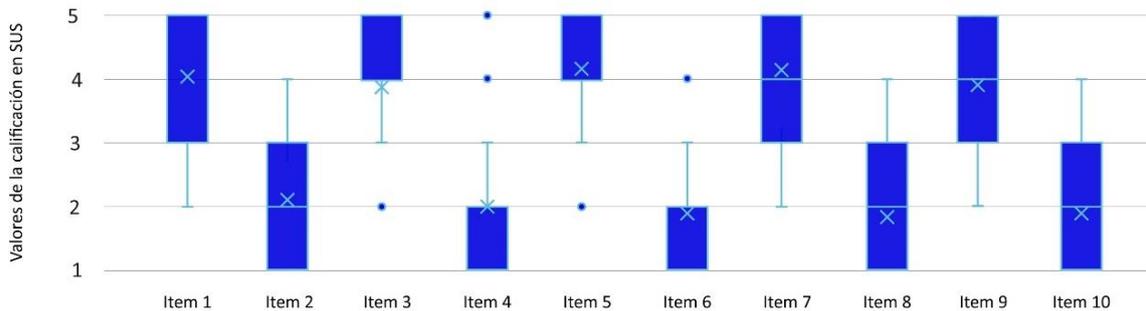


Figura 37: Las puntuaciones obtenidas utilizando el cuestionario SUS para HEUROBOX.

Al examinar detenidamente los datos presentados en la Figura 37, se evidencia que los ítems impares obtuvieron valores promedio de calificación superiores a 3, lo que indica una evaluación positiva de la usabilidad de HEUROBOX. En general, la evaluación de HEUROBOX por parte de las personas expertas fue positiva, como lo demuestra la alta calificación promedio de 4,05 para el ítem "Utilizaría HEUROBOX". Los evaluadores también otorgaron una

calificación promedio de 4 para el ítem "HEUROBOX fue fácil de usar" y 3,9 para el ítem "Me sentí seguro usando HEUROBOX", lo que indica facilidad de uso y confianza en el uso del sistema.

Sin embargo, la evaluación también destacó áreas de preocupación entre las personas expertas con respecto a la complejidad del sistema. Por ejemplo, el ítem "HEUROBOX fue demasiado complejo para mí" obtuvo una calificación promedio de 2,1, lo que indica que algunos participantes encontraron el sistema exigente. El ítem "HEUROBOX fue difícil de usar" también recibió una calificación promedio de 1,85, lo que respalda que algunos participantes tuvieron dificultades para utilizar el sistema. Además, algunos participantes indicaron la necesidad de ayuda para utilizar el sistema, con una calificación promedio de 1,95 para el ítem "Realmente necesito ayuda de alguien para usar HEUROBOX".

En general, los resultados de la evaluación sugieren que HEUROBOX es un sistema usable e intuitivo con margen de mejora en cuanto a complejidad y soporte al usuario. Los hallazgos de los ítems del SUS formulados de manera positiva indican una percepción generalmente positiva de la herramienta entre los evaluadores expertos, con ítems relacionados con la facilidad de uso e integración de varias partes de la herramienta que obtienen puntajes más altos que aquellos relacionados con la complejidad y el soporte al usuario. Estos hallazgos pueden servir como aportes valiosos para informar el diseño y desarrollo de HEUROBOX, con el fin de mejorar su usabilidad y satisfacción de la persona usuaria.

4.4.2 Evaluación AHP

Los códigos de R (Cho, 2019) para ejecutar los algoritmos de AHP (Saaty, 2004) de los datos brutos se pueden encontrar en el conjunto de datos (Apraiz, Mulet Alberola, et al., 2023). Estos códigos de R se presentan en las cuatro secciones principales secuenciadas: (i) preparación del paquete de R, (ii) entradas de datos, (iii) cálculo de los pesos de importancia agregados y (iv) cálculo de las relaciones de consistencia.

La Tabla 36 presenta los resultados del método AHP (Saaty, 2004) aplicado a las respuestas de las personas expertas y calculado en el lenguaje de programación R (Cho, 2019). El análisis revela que la categoría de Seguridad es la más esencial para el diseño de sistemas HRI en entornos industriales. Recibió el peso más alto, 0,469, seguida de Funcionalidad (0,207), Interfaces (0,173) y Ergonomía (0,151). Esto sugiere que, para las personas expertas que participaron en el estudio, es crucial que los sistemas HRI se diseñen y construyan de manera que se minimicen los riesgos de lesiones o daños a las personas usuarias, a los propios robots y a otros elementos del entorno.

Tabla 36: Pesos de importancia agregados en cada grupo de HEUROBOX de acuerdo con AHP.

Categorías	Prioridad	Subcategorías	Prioridad
Seguridad	0,469	Sistemas del robot	0,355
		Planificación del movimiento	0,353
		Medidas organizacionales	0,291
Funcionalidad	0,207	Gestión de errores	0,268
		Tarea	0,225
		Sistema	0,215
		Información	0,152
		Asistencia	0,140
Interfaces	0,173	Visual	0,407

		Voz	0,225
		Gestual	0,195
		Háptica	0,173
Ergonomía	0,151	Ergonomía cognitiva	0,536
		Ergonomía física	0,464

Al desglosar aún más la categoría de Seguridad, se observa que la subcategoría de Sistemas de Robots obtuvo el peso más alto, con 0,355, seguida de Planificación de Movimiento (0,353) y Medidas Organizativas (0,291). Esto sugiere que, para optimizar la seguridad, las consideraciones de diseño para los sistemas de robots deben priorizar estas áreas. Este resultado indica que el diseño de sistemas robóticos seguros requiere centrarse primero en el propio robot, luego en la planificación de su movimiento y en la organización del sistema. Por tanto, la Planificación de Movimiento resalta la importancia de diseñar algoritmos de movimiento eficientes y precisos para los robots. Una planificación adecuada del movimiento es crucial para evitar colisiones, navegar por entornos complejos y garantizar interacciones fluidas y precisas con personas y objetos circundantes. Por último, la categoría de Medidas Organizativas se encuentra en tercer lugar. Aunque las personas expertas le han asignado una importancia relativamente menor en comparación con las otras subcategorías, no se debe subestimar su significado, ya que abarca diversos aspectos relacionados con la organización y gestión del lugar de trabajo. Si bien el enfoque suele centrarse en los aspectos técnicos de la robótica, descuidar los aspectos organizativos puede tener efectos perjudiciales en el rendimiento general y la seguridad del sistema.

En cuanto a la Ergonomía, es posible que se haya clasificado en cuarto lugar en términos de importancia porque, aunque es esencial considerar la HRI desde una perspectiva ergonómica, las demás categorías son más críticas para el rendimiento general y el éxito del sistema robótico. Es importante tener en cuenta que, aunque la ergonomía pueda estar clasificada en una posición inferior en términos de importancia, aún debe recibir una atención adecuada en el diseño y desarrollo de sistemas robóticos. Dentro de la categoría de Ergonomía, la Ergonomía Cognitiva se consideró más importante, con un peso de 0,536, que la Ergonomía Física, con un peso de 0,464. Esto resalta la importancia de diseñar sistemas optimizados para la carga mental y la facilidad de uso. Sugiere que diseñar sistemas robóticos que tengan en cuenta los aspectos cognitivos de la HRI, como la atención, la percepción y la toma de decisiones, es más crítico que diseñar sistemas que se centren únicamente en aspectos físicos, como la postura y el movimiento.

En la categoría de Funcionalidad, la subcategoría de Gestión de Errores recibió el mayor peso, con 0,268, seguida de Tarea (0,225), Sistema (0,215), Información (0,152) y Asistencia (0,140). Esto indica la importancia de desarrollar mecanismos sólidos de manejo de errores y definir tareas relevantes para el sistema. El alto peso asignado a la subcategoría de Tarea indica que las personas expertas dan prioridad a la capacidad del sistema para llevar a cabo tareas específicas de manera precisa y exitosa, resaltando la importancia de diseñar sistemas robóticos optimizados para las tareas específicas que se pretenden realizar en entornos industriales. El peso de la subcategoría de Sistema sugiere que las personas expertas reconocen la importancia de diseñar un sistema bien estructurado e integrado. La inclusión de las subcategorías de Información y Asistencia subraya el valor de proporcionar información relevante y útil a las personas usuarias y ayudar cuando sea necesario. Sin embargo, diseñar sistemas que puedan proporcionar de manera efectiva a las personas usuarias la información y el apoyo necesarios contribuye a una experiencia más intuitiva y amigable, mejorando la funcionalidad general del sistema robótico.

La categoría de Interfaces se ubicó en tercer lugar según las opiniones de las personas expertas. La categoría de Interfaces se considera importante, pero no tan crítica como las categorías de Seguridad y Funcionalidad. Las personas expertas pueden haber priorizado las otras categorías debido a su impacto directo en el éxito y la efectividad general del sistema robótico. La subcategoría de Visual fue considerada la más importante, con un peso de 0,407, seguida de Voz con 0,225, Gestos con 0,195 y Háptica con 0,173. Esto resalta la importancia de desarrollar interfaces que sean visualmente claras y fáciles de entender. Las señales visuales y la retroalimentación desempeñan un papel crucial en facilitar una comunicación e interacción efectivas entre personas y robots. Mediante el diseño de interfaces visualmente atractivas e informativas, las personas usuarias pueden comprender e interpretar fácilmente el estado, las acciones y las intenciones del sistema.

En general, los resultados del análisis AHP proporcionan información valiosa sobre la importancia relativa de las diferentes categorías y subcategorías para el diseño de sistemas robóticos que optimicen la seguridad, la funcionalidad, las interfaces y la ergonomía. Las consideraciones de diseño deben priorizar las áreas identificadas en el análisis para garantizar un sistema exitoso y fácil de usar para las personas usuarias.

El presente estudio obtuvo un Índice de Consistencia (IC) para cada una de las categorías evaluadas. Específicamente, la categoría general exhibió un IC de 0,104, mientras que la categoría de Seguridad mostró un IC de 0,197. En la categoría de Ergonomía, solo se realizó una comparación por pares, lo que resultó en un valor NaN para el IC. Mientras tanto, la categoría de Funcionalidad demostró un IC de 0,129, y la categoría de Interfaces exhibió el IC más bajo de 0,104. Aunque Saaty (2004) establece un límite de no más de 0,1, Ho et al. (2005) señalan que el IC podría relajarse hasta 0,2 para grupos con diferentes niveles de experiencia.

4.5 El indicador de HEUROBOX

Tras la obtención de los coeficientes de importancia mediante el análisis AHP, se ha procedido al cálculo del indicador HEUROBOX. El indicador HEUROBOX se ha diseñado con el propósito de medir y cuantificar el grado de cumplimiento de los principios heurísticos. El cálculo del indicador HEUROBOX se basa en el número de heurísticos cumplidos en cada categoría, teniendo en cuenta los ratios de importancia asignados (Figura 38). Este enfoque proporciona una medida cuantitativa del grado de cumplimiento de los principios heurísticos y permite realizar comparaciones entre diferentes interacciones persona-robot. Al considerar tanto la cantidad de heurísticos cumplidos como la importancia relativa asignada a cada subcategoría, se logra una evaluación más precisa y ponderada de la calidad de la HRI en entornos industriales específicos.

Es importante destacar que este enfoque de cálculo ofrece una visión más detallada y significativa del cumplimiento de los heurísticos, al considerar tanto la cantidad como la relevancia de los mismos. De este modo, se obtiene una evaluación más completa de la calidad de la HRI en entornos industriales, contribuyendo al diseño y desarrollo de sistemas más eficientes y seguros.



Figura 38: Esquema visual de las ponderaciones para calcular el Indicador HEUROBOX.

4.5.1 Identificación de heurísticos cumplidos

En el proceso de evaluación, se lleva a cabo la identificación de los heurísticos cumplidos en cada categoría y subcategoría correspondiente. Esta etapa implica un análisis minucioso del contexto evaluado para determinar qué heurísticos se cumplen en el sistema y requiere la participación de un evaluador experto.

En primer lugar, se realiza un análisis exhaustivo de las características y funcionalidades del sistema robótico. Para ello, se sugiere que el evaluador experto realice una serie de tareas interactivas con el robot, con el objetivo de conocer en profundidad su comportamiento.

Una vez que se ha inspeccionado en profundidad el sistema robótico, se procede a la evaluación heurística. Mediante este análisis detallado, se identifican los heurísticos que se cumplen, es decir, aquellos aspectos del sistema que se ajustan y satisfacen los criterios y directrices establecidos. Estos heurísticos cumplidos representan los aspectos positivos y funcionales de la HRI en el entorno industrial evaluado.

Es importante destacar que la identificación precisa de los heurísticos cumplidos requiere de un conocimiento profundo de los principios heurísticos establecidos y de su aplicación práctica en el contexto específico de estudio. El evaluador experto debe tener experiencia y conocimientos especializados en el campo de la HRI, lo que le permite discernir de manera precisa qué heurísticos se cumplen en la interacción evaluada.

La identificación de los heurísticos cumplidos constituye un paso fundamental en el cálculo del Indicador HEUROBOX, ya que proporciona la base para cuantificar el grado de cumplimiento de los principios heurísticos y, posteriormente, calcular los puntajes parciales y el indicador global. Es responsabilidad del evaluador descartar los heurísticos que no procedan en el caso dado, eliminándolos del número total de heurísticos en esa categoría específica, a fin de garantizar una evaluación precisa y coherente del sistema en estudio.

4.5.2 Cálculo de los puntajes parciales por categoría

La etapa de cálculo de los puntajes parciales se realiza con el propósito de cuantificar el grado de cumplimiento de los heurísticos en cada categoría. Para llevar a cabo este cálculo, se emplea una fórmula basada en la regla de tres, que permite establecer una medida cuantitativa del cumplimiento en relación al total de heurísticos en la categoría correspondiente.

El proceso para calcular el puntaje parcial de una categoría se describe de la siguiente manera:

1. Se multiplica el número de heurísticos cumplidos en la subcategoría por cien. Esta multiplicación tiene como objetivo ampliar la escala de los valores y lograr una representación más significativa y manejable.
2. A continuación, se divide el resultado obtenido en el paso anterior entre el número total de heurísticos presentes en esa categoría (teniendo en cuenta que se han de restar todos aquellos heurísticos que no son aplicables al contexto evaluado). Esta división permite determinar una proporción o porcentaje que refleja el grado de cumplimiento de los heurísticos en la subcategoría en consideración.
3. Posteriormente, se multiplica el resultado obtenido en el paso anterior por el ratio de prioridad correspondiente a la categoría en cuestión. Este factor de prioridad, asignado por expertos, pondera la importancia relativa de cada subcategoría en relación al impacto que tiene en la evaluación de la HRI en entornos industriales.

La fórmula utilizada para el cálculo del puntaje parcial de una subcategoría se expresa de la siguiente manera:

$$\text{Puntaje parcial de la categoría} = \frac{nc \cdot 100}{nt} \cdot r \quad (15)$$

Siendo,

nc= Número de heurísticos cumplidos en la categoría

nt= Número de heurísticos totales en la categoría

r= Ratio de importancia de la categoría

4.5.3 Cálculo del indicador HEUROBOX

El cálculo del indicador HEUROBOX consiste en la suma de los puntajes parciales obtenidos en todas las categorías evaluadas. Esta suma permite obtener un valor total que representa el grado de cumplimiento y calidad de la HRI en entornos industriales, considerando todos los aspectos evaluados mediante los principios heurísticos y los factores de ponderación asignados.

La fórmula utilizada para calcular el indicador HEUROBOX se expresa de la siguiente manera:

$$\text{Indicador HEUROBOX} = \sum \text{Puntaje parcial de la categoría} \quad (16)$$

Al sumar los puntajes parciales de todas las categorías, se obtiene un valor total que proporciona una medida global y cuantitativa del desempeño de la HRI en entornos industriales. Este indicador HEUROBOX refleja la sinergia de los diferentes aspectos evaluados, abarcando la seguridad, la ergonomía, la funcionalidad y las interfaces.

El valor resultante del indicador HEUROBOX permite comparar y evaluar diferentes configuraciones de HRI, identificar áreas de mejora y tomar decisiones informadas para optimizar la interacción en entornos industriales. Además, al considerar los puntajes parciales ponderados por la importancia asignada a cada categoría, se asegura una evaluación equilibrada y completa de la HRI.

4.6 Discusión y conclusiones

En este apartado, se recoge una discusión detallada sobre HEUROBOX en tres aspectos: (i) su adecuación a los requisitos previamente establecidos, (ii) los resultados obtenidos en la evaluación de usabilidad y, por último, (iii) las fortalezas como las limitaciones de esta herramienta.

En la literatura actual, existe una brecha en la evaluación heurística de la HRI en entornos industriales desde una perspectiva de UX y HCD. Se considera que falta un enfoque holístico y centrado en las personas, ya que hasta ahora la literatura ha estado en gran medida centrada en el robot (Prati et al., 2021). Con el objetivo de abordar esta limitación, se propone un nuevo instrumento basado en la evaluación heurística llamado HEUROBOX para evaluar los sistemas HRI y, así, mejorar la UX, la aceptación de la tecnología y el bienestar general. La herramienta propuesta, HEUROBOX, aborda esta brecha al recopilar heurísticas propuestas por diversos expertos en diferentes ámbitos. Como resultado, proporciona una herramienta integral que incorpora múltiples perspectivas.

HEUROBOX es una novedosa herramienta de evaluación heurística propuesta para evaluar el grado de alineación entre el robot y las características contextuales con las necesidades humanas, a fin de facilitar interacciones eficientes y satisfactorias en entornos industriales. La herramienta consta de cuatro categorías: Seguridad, Ergonomía, Funcionalidad e Interfaces. El nivel básico de HEUROBOX abarca aspectos fundamentales de la evaluación de HRI e incorpora 84 principios heurísticos. Por otro lado, el nivel avanzado comprende una variedad de elementos o funciones específicas e incluye 228 principios heurísticos, incluidos los del nivel básico. Es importante destacar que evaluar cada categoría es crucial para garantizar de manera integral una experiencia fluida y satisfactoria.

4.6.1 La adecuación de HEUROBOX a los requisitos

HEUROBOX, como una herramienta integral de evaluación heurística, ha sido diseñada específicamente para cumplir con los requisitos establecidos para evaluar la HRI en entornos industriales. En este sentido, la Tabla 37 resume cómo HEUROBOX se ajusta a los requisitos establecidos.

Tabla 37: La adecuación de HEUROBOX a los requisitos establecidos.

Requisitos	Cómo HEUROBOX cumple con ellos
Seguridad, confianza y seguridad percibida	HEUROBOX está diseñado con la seguridad como máxima prioridad. Incorpora una categoría específica de Seguridad (con un total de 30 heurísticos) y características avanzadas de seguridad, como detección y evitación de colisiones, para garantizar la seguridad de todas las personas involucradas en la HRI. Además, HEUROBOX está diseñado para establecer la confianza entre las personas operarias humanas y los robots, asegurando que el robot protegerá los intereses y el bienestar de todos los individuos en el equipo.
Ergonomía física	HEUROBOX ha sido diseñado para abordar la ergonomía física en el lugar de trabajo, garantizando que las personas trabajadoras estén cómodas y no corran riesgo de trastornos musculoesqueléticos. De hecho, contiene una subcategoría específica sobre ergonomía física que consta de un total de 4 heurísticos en la evaluación básica y 25 en la evaluación avanzada. El sistema puede ayudar a prevenir la incomodidad de la persona trabajadora y el agravamiento de condiciones posturales o físicas en el entorno de trabajo.
Ergonomía cognitiva y emociones	HEUROBOX está diseñado teniendo en cuenta la ergonomía cognitiva y las emociones. De hecho, tiene una subcategoría específica sobre ergonomía cognitiva que consta de un total de 9 heurísticos en la evaluación básica y 20 en la evaluación avanzada. Las diferentes categorías de HEUROBOX ayudarían a mejorar aspectos como la carga mental de trabajo, el estrés físico y mental, la facilidad de aprendizaje, la utilidad, la eficiencia, el uso intuitivo, la confiabilidad, la estimulación y la confianza.

Inclusividad	HEUROBOX prioriza la inclusión, reconociendo que no todas las personas usuarias pueden acceder a los robots debido a características diversas como limitaciones visuales, auditivas, motoras o cognitivas. El sistema está diseñado teniendo en cuenta la accesibilidad, evitando la creación de barreras que impidan a las personas con discapacidades utilizar el sistema robótico. HEUROBOX implementa las pautas de accesibilidad de Qbilat et al. (2021) a través de las diferentes categorías.
Multimodalidad	HEUROBOX ofrece una variedad de interfaces para permitir la comunicación multimodal, como comunicación gráfica, basada en voz o basada en gestos. Incluye pantallas visuales, gestos, habla y lenguaje natural, e interacciones físicas y hápticas. Esta variedad de interfaces permite a las personas usuarias interactuar con el sistema de la manera que mejor se adapte a sus necesidades.
Tipo de robot y funcionalidades	HEUROBOX está diseñado para adaptarse a diferentes tipos de robots y funcionalidades, lo que permite flexibilidad en el lugar de trabajo. Está diseñado para admitir diferentes tipos de <i>software</i> , lo que le permite realizar diferentes funcionalidades según sea necesario.

4.6.2 Evaluación sobre HEUROBOX

La evaluación de usabilidad a través de las valoraciones de expertos de la herramienta HEUROBOX ha obtenido resultados favorables, lo que indica la efectividad de la herramienta para evaluar la calidad de la HRI en entornos industriales. Los resultados positivos de la evaluación de usabilidad de HEUROBOX indican que es un enfoque prometedor para garantizar que las interacciones persona-robot sean eficientes y satisfactorias. Además, el índice general de 76 obtenido de la evaluación de usabilidad destaca el potencial de HEUROBOX para proporcionar una evaluación integral de la calidad de la HRI, considerando diversos factores como la seguridad, la ergonomía, la funcionalidad y las interfaces. Estos hallazgos respaldan el uso de HEUROBOX como una herramienta valiosa para evaluar y mejorar las interacciones persona-robot en entornos industriales desde una perspectiva centrada en el ser humano.

Los hallazgos mencionados anteriormente proporcionan información valiosa para los diseñadores y desarrolladores al priorizar sus esfuerzos de diseño para optimizar aspectos críticos de la HRI. Los diseñadores pueden aprovechar estos resultados como guía para centrarse en áreas clave que han sido identificadas como importantes por expertos en el campo. Por ejemplo, las respuestas de las personas expertas, obtenidas mediante la aplicación del AHP, indican que la seguridad es la categoría general más crítica, seguida de la funcionalidad, las interfaces y la ergonomía. El enfoque en la seguridad es especialmente importante en sistemas robóticos que interactúan con personas, como los robots colaborativos, ya que se requiere que estos sistemas sean seguros y confiables para trabajar junto a las personas. No obstante, la seguridad es uno de los aspectos más estudiados de la robótica industrial y existen numerosas normas con las que los robots deben cumplir para garantizar los aspectos de seguridad (ISO 10218, 2011).

El peso elevado asignado a la subcategoría de Sistemas de Robot dentro de la categoría de Seguridad enfatiza la importancia de considerar al propio robot al diseñar sistemas robóticos seguros. Este conocimiento sugiere que los diseñadores deben prestar atención cuidadosa a la construcción del robot, su funcionalidad y sus características de seguridad inherentes para minimizar los riesgos y garantizar el bienestar de la persona usuaria. De manera similar, la priorización del manejo de errores resalta la necesidad de mecanismos robustos para manejar errores o fallos que puedan ocurrir durante la operación del sistema. Mediante la implementación de estrategias efectivas de manejo de errores, como la detección de errores, la recuperación y la tolerancia a fallos, se puede mejorar la confiabilidad y el rendimiento del sistema robótico. Asimismo, el énfasis en la subcategoría Visual implica que los diseñadores deben priorizar la creación de interfaces visualmente atractivas e intuitivas.

Este conocimiento enfatiza la importancia de señales visuales claras e informativas para facilitar la comunicación y la interacción efectiva entre las personas y los robots.

Al utilizar estos conocimientos, los diseñadores y desarrolladores pueden asignar estratégicamente sus recursos y esfuerzos, centrándose en los aspectos críticos identificados por las personas expertas. Este enfoque garantiza que la HRI esté optimizada en términos de seguridad, funcionalidad, ergonomía e interfaces, lo que conduce a una mejor UX, una mayor aceptación por parte de las personas usuarias y un bienestar general. En general, los resultados discutidos en este estudio proporcionan orientación práctica para los diseñadores y desarrolladores, permitiéndoles priorizar los esfuerzos de diseño y optimizar sus sistemas en aspectos críticos de la HRI.

4.6.3 Fortalezas y limitaciones de HEUROBOX

En el campo de la HRI en entornos industriales, la evaluación de la UX y la aceptación de la tecnología de los sistemas es crucial. En este sentido, HEUROBOX ofrece un enfoque estructurado para evaluar la HRI, permitiendo a personas investigadoras, desarrolladoras o profesionales del sector identificar y evaluar de manera sistemática la efectividad de diferentes aspectos de la interacción de manera rápida y económica.

HEUROBOX proporciona un marco para evaluar sistemas robóticos que es tanto integral como sistemático. El uso de este conjunto de heurísticas permite identificar una amplia gama de posibles problemas de usabilidad que de otra manera podrían pasarse por alto. Además, la naturaleza sistemática de HEUROBOX garantiza que se consideren todos los aspectos de la HRI, incluyendo tanto los factores tecnológicos como las personas.

Las heurísticas pueden proporcionar conocimientos valiosos sobre la UX de la colaboración, ayudando a comprender cómo las personas perciben e interactúan con los robots. Como otras herramientas de evaluación heurística, HEUROBOX puede utilizarse de manera iterativa a lo largo del proceso de diseño (Nielsen & Molich, 1990; Rubin & Chisnell, 2008). Además, HEUROBOX puede utilizarse en conjunto con otros métodos de evaluación para una evaluación más completa. La combinación de diferentes métodos de evaluación puede ayudar a proporcionar una perspectiva más completa e identificar posibles problemas que podrían haberse pasado por alto utilizando solo heurísticas. La lista propuesta de heurísticas también es escalable, lo que permite la adición de nuevas heurísticas que reflejen y evalúen nuevas funcionalidades y diseños de sistemas robóticos que afecten a las interacciones, así como aspectos novedosos del fenómeno detrás de la HRI que deben ser evaluados. Puede utilizarse en conjunto con otros métodos de evaluación para una evaluación más completa. La combinación de diferentes métodos de evaluación puede ayudar a proporcionar una perspectiva más completa e identificar posibles problemas que podrían haberse pasado por alto utilizando solo heurísticas (Apraiz, Lasa, & Mazmela, 2023).

Si bien HEUROBOX puede proporcionar una evaluación precisa de la HRI, también tiene la limitación de que el evaluador debe tener un buen conocimiento del robot específico que se está estudiando. Esto se debe a que HEUROBOX incluye un conjunto de heurísticas detalladas que requieren un conocimiento preciso de las características, funciones y capacidades del robot en cuestión. Sin embargo, esta limitación también puede considerarse una fortaleza, ya que permite a los investigadores, desarrolladores o profesionales llevar a cabo una evaluación muy específica y dirigida de la HRI. Las evaluaciones heurísticas a veces pueden ser subjetivas y depender de la experiencia y conocimientos del evaluador, lo que puede introducir sesgos y variabilidad en los resultados. Además, se puede percibir como un proceso que requiere mucho tiempo y trabajo intensivo para los evaluadores.

En cuanto a las interacciones multimodales, que implican la integración de múltiples modalidades sensoriales para una comunicación más natural con los robots, tienen el potencial

de crear interfaces de usuario y canales de interacción poderosos. Si bien la lista propuesta de heurísticas es aplicable a diferentes tipos de interfaces y, por lo tanto, a entornos multimodales, puede haber superposición entre las heurísticas propuestas dentro de las subcategorías de la sección de Interfaces. Esta superposición podría dar lugar a una situación en la que un robot con dos tipos de interfaces parezca cumplir un menor porcentaje de heurísticas. Sin embargo, cuando estas dos interfaces se complementan entre sí de manera efectiva, el nivel general de interacción puede mejorar en consecuencia. Por lo tanto, es importante tener en cuenta que la evaluación de interfaces en un contexto multimodal requiere un enfoque matizado. La superposición entre las heurísticas debe examinarse cuidadosamente, teniendo en cuenta la interacción entre diferentes modalidades y cómo contribuyen a la UX general.

A pesar de estas limitaciones, HEUROBOX tiene el potencial de proporcionar una evaluación precisa de la HRI. Su conjunto detallado de heurísticas requiere que los evaluadores tengan un buen entendimiento del robot específico que se está estudiando, pero esta limitación también puede considerarse una fortaleza, ya que permite una evaluación específica y dirigida de la HRI. Al utilizar HEUROBOX en conjunto con otros métodos de evaluación, los evaluadores pueden obtener una comprensión integral de las fortalezas y debilidades del robot en cuestión. Además, al utilizar HEUROBOX para evaluar diferentes robots o diferentes versiones del mismo robot, los evaluadores pueden comparar los resultados en diferentes escenarios e identificar patrones o tendencias a lo largo del tiempo, proporcionando conocimientos valiosos sobre el progreso del desarrollo de robots.

Principales conclusiones del capítulo

- Se presenta **HEUROBOX, una herramienta de evaluación basada en heurísticas desarrollada y validada para la HRI en entornos industriales** y que se basa en cuatro categorías diferentes: (i) la **seguridad**, (ii) la **funcionalidad**, (iii) la **ergonomía** y (iv) las **interfaces**. En el Anexo B se presenta el listado completo.
- HEUROBOX simplifica la evaluación de la HRI y permite identificar áreas de mejora en el sistema en términos de UX. No obstante, para obtener una evaluación completa y precisa de la UX, es recomendable complementarla con otras técnicas de evaluación, como las pruebas de usabilidad, tal y como se propone en el modelo ITPX-ROBOTS.
- La **seguridad** juega un papel fundamental en la HRI. En este contexto, HEUROBOX se enfoca en cuestiones de diseño que tienen como objetivo evitar cualquier tipo de perjuicio, tanto a las personas como a los propios robots. Esto involucra la implementación de medidas de seguridad, tales como sistemas de parada de emergencia, barreras de protección y sistemas de detección de colisiones. Además, al diseñar el robot, es crucial tener en cuenta las limitaciones físicas y cognitivas de las personas que lo utilizan, así como la capacidad de adaptación a situaciones inesperadas.
- La **ergonomía** se enfoca en la interacción entre la manera en que las personas y los robots interactúan, considerando aspectos como la postura, los movimientos y la comodidad. Al diseñar el robot, es esencial asegurarse de que las personas que lo utilizan puedan hacerlo sin experimentar tensiones físicas innecesarias ni molestias. Además, HEUROBOX también tiene en cuenta aspectos que permiten mejorar la

ergonomía cognitiva.

- La **funcionalidad** representa un componente esencial en la planificación del diseño del robot, ya que establece la habilidad del robot para llevar a cabo sus tareas previstas de manera eficiente y efectiva. Es esencial que la funcionalidad del robot se ajuste al contexto de uso específico y tenga la capacidad de lidiar con cualquier situación imprevista que pueda presentarse.
- Las **interfaces** son una categoría crucial en el diseño de la HRI, ya que establecen los medios de comunicación entre la persona usuaria y el robot. Estas interfaces engloban diversas consideraciones de diseño, que incluyen aspectos visuales, vocales, gestuales y táctiles. Las interfaces visuales deben ser diseñadas de manera que presenten la información necesaria de manera clara y organizada, haciendo uso de gráficos e iconos sencillos. Las interfaces de voz deben ser diseñadas de forma natural y parecida a las conversaciones humanas, manteniendo una estructura coherente en el diálogo entre la persona y el sistema. En cuanto a las interfaces de gestos, deben ser intuitivas y naturales para que la persona las utilice con facilidad, mientras que las interfaces hápticas deben proporcionar retroalimentación táctil suficiente al usuario.
- A pesar de que las cuatro categorías que integran HEUROBOX son distintas, están intrínsecamente interconectadas y cada una tiene un impacto significativo en las demás. Por ejemplo, la seguridad guarda una relación estrecha con la ergonomía, ya que un robot diseñado de manera ergonómica disminuirá las posibilidades de ocasionar daño al usuario. Además, la funcionalidad está íntimamente ligada a las interfaces, dado que estas deben ser configuradas de manera que permitan al robot llevar a cabo sus tareas de forma eficiente y efectiva. Por último, la UX y la aceptación de la tecnología se verán influidas por las cuatro categorías, puesto que la seguridad, la ergonomía, la funcionalidad y las interfaces del robot incidirán en la percepción global de la interacción.
- Dependiendo de la complejidad inherente del sistema robótico, el conjunto de heurísticas que debe cumplir puede variar para lograr una experiencia satisfactoria. En otras palabras, un sistema más sencillo podría satisfacer únicamente el nivel básico de heurísticas y aún así proporcionar una experiencia gratificante, mientras que un sistema más complejo podría cumplir también con un mayor nivel de heurísticas, pero seguir generando una experiencia frustrante para la persona usuaria. Es por ello, por lo que se ha propuesto la inclusión del Nivel Avanzado, que permite a una persona experta elegir el conjunto de heurísticas más apropiado para el robot en cuestión. Sin embargo, se sugiere llevar a cabo una evaluación completa de la UX en la HRI. Esto implica considerar de manera tanto objetiva como subjetiva las características de la persona que utiliza el robot, el contexto en el que se lleva a cabo la interacción y el sistema en sí mismo, antes, durante y después del encuentro entre persona y robot.

Capítulo V

Desarrollo y validación del cuestionario Human-Robot Experience (HUROX)

5 DESARROLLO Y VALIDACIÓN DEL CUESTIONARIO HUMAN-ROBOT EXPERIENCE (HUROX)

Los cuestionarios son un instrumento de uso común que proporcionan información cuantitativa sobre la percepción de la persona usuaria. Sin embargo, aún hay pocas aplicaciones en el campo de la HRI en contextos industriales, pero demuestran que los cuestionarios de UX presentan una gran oportunidad para evaluar. En gran medida, porque recopilan constructos que pueden ser interesantes de aplicar en el contexto de la Industria 5.0 y así obtener información sobre la percepción de la persona usuaria. Al proponer herramientas ágiles, como cuestionarios, para evaluar la interacción de la persona operaria con las máquinas, es posible identificar los factores que deben mantenerse o mejorarse para optimizar esta interacción y, así, mejorar los procesos que deben ocurrir en las fábricas inteligentes.

La integración de cuestionarios con herramientas objetivas o cualitativas puede proporcionar una comprensión más completa tanto de la aceptación de la tecnología como de la UX. Aunque existe una amplia literatura sobre el uso de cuestionarios para evaluar la aceptación de la tecnología y la UX, no existe un cuestionario para evaluar la UX y la aceptación de la tecnología en un contexto colaborativo industrial. Por lo tanto, crear un cuestionario para la persona usuaria final para evaluar la UX en la HRI es un paso crucial para mejorar el diseño y la aceptación de los sistemas HRI en contextos industriales. Al obtener retroalimentación sistemática y estructurada de las personas usuarias finales, los investigadores y profesionales pueden identificar problemas y mejorar la UX en general, lo que conduce a sistemas HRI más efectivos, eficientes y satisfactorios. Además, un cuestionario para la persona usuaria final puede ayudar a comprender las actitudes de las personas usuarias finales hacia la tecnología e identificar áreas potenciales de mejora en el sistema HRI.

El cuestionario Human-Robot Experience (HUROX) se desarrolla con el propósito de enriquecer el modelo ITPX-ROBOTS con un carácter perceptual y considerando los aspectos relevantes para la HRI identificados en el marco teórico HRIX. Este cuestionario se integra en la tercera fase del modelo ITPX-ROBOTS, donde se recopila la percepción del usuario tras la interacción con el robot.

Esta sección presenta el desarrollo y la validación psicométrica del cuestionario HUROX, que es un cuestionario para la persona usuaria final que evalúa la aceptación de la tecnología y la UX a través de 41 ítems en 9 constructos (utilidad percibida, facilidad de uso percibida, seguridad percibida, comodidad, facilidad de aprendizaje, controlabilidad, actitud, confianza y satisfacción).

5.1 Metodología

En el presente estudio, la elaboración del cuestionario y la evaluación psicométrica se realizaron en cuatro etapas, según el método propuesto por Jahangiri et al. (2021). Se explican las siguientes etapas (Figura 39):

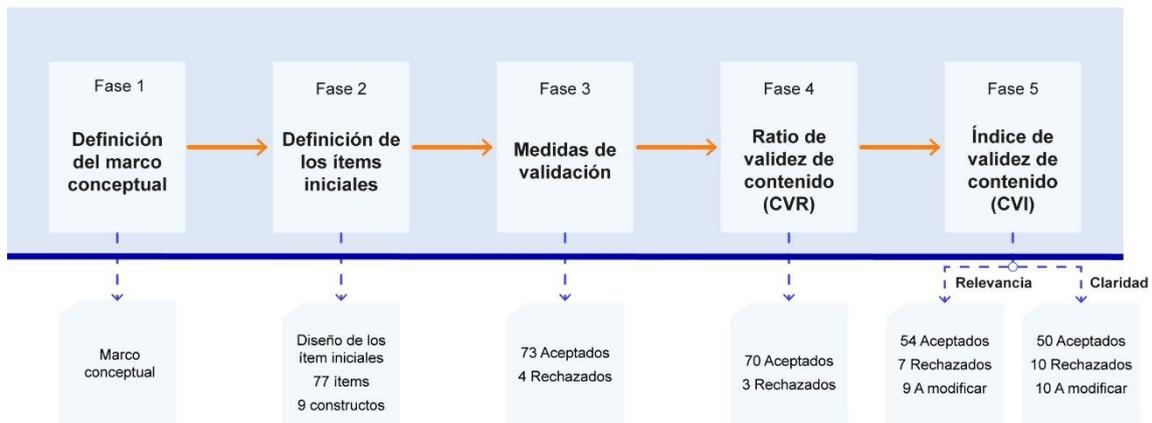


Figura 39: Metodología seguida para el desarrollo del nuevo cuestionario HUROY.

La creación, desarrollo y validación del cuestionario HUROY se llevó a cabo en inglés a pesar de que la presente documentación haya sido traducida al castellano. Se adoptó este enfoque para garantizar la consistencia y precisión en la comunicación de los constructos y los ítems. La creación del cuestionario implicó un riguroso proceso de desarrollo y adaptación de las escalas y los ítems existentes de acuerdo con los objetivos y el contexto específico de este estudio. Se realizaron exhaustivas revisiones bibliográficas y se consultaron fuentes académicas relevantes para identificar y seleccionar constructos pertinentes y sus respectivas definiciones. Posteriormente, se elaboraron los ítems que componen cada uno de los constructos, siguiendo pautas de redacción claras y coherentes. La validación del cuestionario se llevó a cabo mediante un enfoque metodológico sólido y estandarizado y las personas participantes completaron el cuestionario en su versión original en inglés.

5.1.1 Fase 1: Definición del marco conceptual

Partiendo del marco conceptual HRIX que se ha presentado en el Capítulo III y que integra aspectos experienciales e instrumentales, el presente estudio aborda las características del robot, la persona y el contexto. Por lo tanto, el marco conceptual de este estudio sostiene que la interacción entre un robot y una persona está determinada por una combinación de las características del robot, las características de la persona usuaria y los factores contextuales. Estas características de la interacción impactan en los componentes de la UX, que tal y como se muestran en el modelo HRIX, son: (i) la utilidad, (ii) la facilidad de uso, (iii) la ergonomía, (iv) la aprendibilidad y (v) la controlabilidad. Estos, a su vez, tienen consecuencias en los factores humanos: (i) actitud e intención de uso, (ii) confianza, (iii) seguridad percibida, y (iv) satisfacción.

5.1.2 Fase 2: Definición de los ítems iniciales

En la segunda fase, en base a la bibliografía, se diseñaron los ítems para los constructos del marco conceptual. Para comenzar el diseño de los ítems iniciales, se reunió una colección considerable de ítems potenciales para su evaluación. Esta colección se obtuvo de cuestionarios establecidos (Bartneck et al., 2008; Brooke, 1996; Lund, 2001; Schmidtler et al., 2017; von der Pütten & Bock, 2018), estudios internos previos (Apraiz et al., 2022; Apraiz & Lasa, 2020; Aranburu et al., 2020; Mazmela et al., 2019, 2020), y sesiones de brainstorming. Posteriormente, se llevó a cabo un proceso de curación para eliminar o reformular los ítems que no eran adecuados para su aplicación en el dominio de la HRI.

5.1.2.1 Fuente de los ítems

Tras establecer los constructos que formarían el cuestionario, el conjunto inicial de ítems se extrajo de diversas fuentes de los campos de la psicología, la interacción persona-ordenador, la HRI, la experiencia de la persona usuaria y el modelo de aceptación. Se trata de cuestionarios que han sido validados o utilizados en experimentos. Las fuentes utilizadas incluyen los cuestionarios indicados en la Tabla 38.

Tabla 38: Fuente de los ítems para el cuestionario HUOX.

Cuestionario	Referencia	Validado	Constructos	Construcción de los ítems	N.º de ítems
SUS	Brooke (1996)	Sí	Usabilidad y aprendibilidad	Escala de Likert	10
USE	Lund (2001)	Sí	Aprendibilidad, utilidad y facilidad de uso	Escala de Likert	30
Godspeed	Bartneck et al. (2008)	Sí	Antropomorfismo, animadversión, simpatía, inteligencia percibida y seguridad percibida	Pares semánticos	24
Autogenerado	Daniel et al. (2013)	No	Confianza en la automatización	Escala de Likert	19
Autogenerado	Lasota & Shah (2015)	No	Satisfacción con un robot como compañero de equipo	Escala de Likert	8
QUEAD	Schmidtler et al. (2017)	Sí	Utilidad percibida, facilidad de uso percibida, emociones, actitud y comodidad.	Escala de Likert	16
UEQ	Hinderks et al. (2018)	Sí	Atractividad, perspicacia, eficacia, fiabilidad, estímulo y novedad.	Pares semánticos	26

Además, el nuevo cuestionario también se basa en los siguientes marcos conceptuales:

- El estudio de Berg et al. (2008), que resume las principales conclusiones del trabajo sobre las motivaciones y necesidades resultados individuales, experiencias positivas, experiencias negativas involuntarias y resiliencia.
- El modelo HRCAM desarrollado por Bröhl et al. (2019), que define los elementos clave de la percepción humana en HRI.

La Tabla 39 muestra los constructos que componen la versión inicial del cuestionario, incluyendo sus respectivas definiciones y el número total de ítems iniciales para cada constructo que se originaron a partir de los cuestionarios y marcos conceptuales mencionados.

Tabla 39: Constructos iniciales y su definición del cuestionario inicial.

Constructo	Definición	N.º de ítems
Utilidad percibida	El grado en que una persona cree que el uso de un sistema en particular mejoraría su desempeño laboral (Davis, 1989).	12
Facilidad de uso percibida	El grado en que una persona cree que el uso de un sistema en particular sería libre de esfuerzo y fácil de aprender (Davis, 1989).	16
Seguridad percibida	El grado en que una persona percibe que un sistema en particular puede adquirir nuevos conocimientos o habilidades (Akalin et al., 2022).	7
Ergonomía	El grado en que una tecnología se adapta a las habilidades y limitaciones físicas y cognitivas de la persona usuaria, incluyendo aspectos como comodidad, facilidad de uso y accesibilidad (Moroney & Lilienthal, 2008).	3
Aprendibilidad	El grado en que las personas usuarias pueden realizar con éxito una tarea	5

cuando se encuentran con una interfaz por primera vez, así como el grado en que pueden volverse competentes en esa tarea con repeticiones crecientes de uso (Joyce, 2019).

Controlabilidad	El grado en que una persona se siente en control de una tecnología y sus acciones, y la capacidad para modificar su comportamiento según sus necesidades o preferencias (Sarker & Wells, 2003).	11
Actitud	La evaluación general positiva o negativa de un usuario hacia una tecnología, componentes afectivos y cognitivos. Es un predictor de la intención de uso de la persona usuaria (Davis, 1989; Fishbein & Ajzen, 1977).	7
Confianza	El grado en que una persona cree que una tecnología funcionará según lo previsto y no les causará daño a ellos ni a otros (Mayer et al., 1995).	3
Satisfacción	El grado en que una persona está satisfecha con el uso de una tecnología, como resultado de la utilidad, facilidad de uso y confianza en su conjunto (K. C. Lee & Chung, 2009).	13
TOTAL		77

5.1.2.2 Construcción de los ítems

De acuerdo con Lewis & Mayes (2014), existen dos enfoques básicos para la construcción de los ítems para cuestionarios estandarizados diseñados para evaluar la UX. El primero consiste en construir los ítems de tipo conformidad (las personas participantes indican en qué medida están de acuerdo con una determinada afirmación) como QUEAD (Schmidtler et al., 2017), USE (Lund, 2001), SUS (Brooke, 1996). La segunda consiste en construir pares semánticos (las personas participantes posicionan sus valoraciones entre extremos opuestos) como el cuestionario Godspeed (Bartneck et al., 2008), el PANAS (Watson et al., 1988), AttrakDiff (Hassenzahl, 2001) o el UEQ (Hinderks et al., 2018).

De acuerdo con la práctica habitual en la investigación con cuestionarios, los ítems se formularon de tipo conformidad con escalas de valoración de siete puntos basadas en el formato Likert ya que, según Preston & Colman (2000), las puntuaciones más fiables se derivaban de esta escala. La escala de Likert es un enfoque ampliamente adoptado en la investigación en ciencias sociales que permite cuantificar actitudes o percepciones subjetivas pidiendo a los encuestados que indiquen su nivel de acuerdo o desacuerdo con una afirmación utilizando una escala que va desde totalmente en desacuerdo hasta totalmente de acuerdo (Allen & Seaman, 2007). Al utilizar el formato de la escala Likert, los ítems se diseñaron para obtener respuestas matizadas y cuantificables de las personas participantes, lo que permite un análisis detallado de sus percepciones y actitudes hacia el sistema robótico investigado.

La versión inicial del cuestionario constaba de 77 ítems organizados en 9 constructos. Los ítems iniciales se presentan en la Tabla 41.

5.1.3 Fase 3: Medidas de validación

En la tercera fase se evaluó la validez del cuestionario tras la formulación inicial del contenido de los ítems. La validez de un instrumento viene determinada por su capacidad para medir con precisión el constructo que se pretendía evaluar (Jahangiri et al., 2021). La validez aparente de cada ítem se evaluó utilizando la Puntuación de Impacto del Ítem (IIS, por sus siglas en inglés), según la cual sólo se consideraron aceptables los ítems con una puntuación IIS de 1,5 o superior. Se invitó a un panel de 15 expertos a evaluar la validez aparente del cuestionario mediante calificaciones de la importancia de cada ítem utilizando una escala Likert de cinco puntos que iba desde "nada importante" (1), "algo importante" (2), "moderadamente importante" (3), "importante" (4), "muy importante" (5).

La puntuación IIS de cada ítem se calculó como el producto del porcentaje de expertos que asignaron una puntuación de 4 o 5 al ítem y la puntuación de valoración media de cada ítem, según la fórmula 1:

$$IIS = \text{porcentaje de expertos que asignaron una puntuación de 4 o 5 x} \\ \text{puntuación media de cada ítem} \quad (17)$$

5.1.4 Fase 4: Ratio de validez del contenido (CVR)

El ratio de validez de contenido (CVR, por sus siglas en inglés) depende del análisis lógico, y su determinación se basa en un juicio subjetivo e individual. Para medir este índice, se pidió a las personas expertas que clasificaran cada ítem basándose en una escala Likert de tres puntos (1=innecesario, 2=útil, pero no esencial, 3=necesario). El CVR propuesto por (Lawshe, 1975) es una transformación lineal de un nivel de acuerdo proporcional sobre cuántos "expertos" dentro de un panel califican un ítem como "esencial" o "útil" calculado de la siguiente manera (Ayre & Scally, 2014):

$$CVR = \frac{ne - \frac{n}{2}}{\frac{n}{2}} \quad (18)$$

Donde,

n= el número total de expertos

ne= número de expertos que eligieron la opción "necesario" o "útil, pero no esencial" para el ítem del cuestionario.

En función del número de expertos que evalúan los ítems, el valor mínimo aceptable del CVR debe basarse en los valores introducidos por Lawshe (1975) para la adecuación de la validez de contenido. En el presente estudio participaron 15 expertos; por lo tanto, si el valor CVR del ítem era igual o superior a 0,49, el ítem era aceptado.

5.1.5 Fase 5: Índice de validez del contenido (CVI)

Según Yusoff (2019), establecer la validez de contenido es vital para respaldar la validez de una herramienta de evaluación como los cuestionarios, especialmente para fines de investigación. Para medir el índice de validez del contenido (CVI, por sus siglas en inglés), primero se pidió a las personas expertas que calificaran los ítems en términos de Relevancia desde "nada relevante" (1), "relativamente relevante" (2), "relevante" (3), hasta "completamente relevante" (4). A continuación, las personas expertas debían puntuar los ítems en términos de claridad, desde "nada claro" (1), "relativamente claro" (2), "claro" (3), hasta "muy claro" (4). Según Lawshe (1975), el CVI es un cálculo estadístico útil para rechazar o retener ítems específicos.

Para cada ítem, el I-CVI puede calcularse contando el número de expertos que calificaron el ítem como 3 o 4 y dividiendo ese número por el número total de expertos, es decir, la proporción de acuerdo sobre la validez de contenido de un ítem.

$$I - CVI = \frac{ne}{n} \quad (19)$$

Donde,

n= número total de expertos.

ne= número de expertos que eligieron la opción “3” o “4” para el ítem del cuestionario.

5.1.6 Caracterización de las personas expertas

La evaluación fue efectuada por 15 expertos que demostraron una formación profesional y áreas de especialización diversas (Tabla 40). El 60% de las personas expertas eran hombres y su experiencia profesional oscilaba entre 2 y 24 años. La experiencia de las personas participantes abarcaba una amplia gama de ámbitos, como la experiencia de la persona usuaria, la HRI, el diseño centrado en el ser humano, la ergonomía, la inclusión digital, el diseño de interacción y la fabricación. Esta amplitud de conocimientos permitió a las personas participantes aportar valiosas ideas para el desarrollo del cuestionario y garantizar su pertinencia para los posibles usuarios finales en diversos contextos industriales.

Tabla 40: Caracterización de las personas expertas que participaron en la validación del cuestionario HUROY.

Id	Género	Conocimientos	Años en el campo
1	Hombre	User Experience	4
2	Mujer	User Experience in Human-Robot Interaction	4
3	Hombre	User Experience in game-based systems	2
4	Hombre	Research scholar	4
5	Hombre	User Experience and interaction	4
6	Hombre	Computer science	15
7	Hombre	User Experience and Ergonomics	7
8	Hombre	User Experience and Ergonomics	24
9	Mujer	Human-centred design	3
10	Mujer	Digital inclusion	3
11	Mujer	UX and Technology Acceptance in industrial environments	7
12	Mujer	UX and Strategic Design	4
13	Hombre	Operations Excellence, Manufacturing, Human centred design	8
14	Mujer	Human-centred design	3
15	Hombre	Interaction design	12

5.2 Resultados

En esta sección se presentan el diseño y los resultados de la evaluación psicométrica del cuestionario HUROY.

5.2.1 Validez aparente del cuestionario

El inicio de la investigación, el cuestionario constaba de 77 ítems. Para evaluar su validez aparente, se utilizó el IIS como herramienta de evaluación. Está ampliamente reconocido que un IIS no inferior a 1,5 se considera aceptable para que un ítem de investigación demuestre validez aparente (Jahangiri et al., 2021). Por lo tanto, sólo los ítems con un IIS superior a 1,5 se consideraron aceptables para su inclusión en el estudio. Los resultados de la evaluación de la validez aparente, presentados en la Tabla 41, detallan el IIS atribuido a cada ítem. Como se ha indicado, se rechazaron 3 ítems, PEU6, PEU11 y C8.

5.2.2 Ratio de Validez de Contenido del Cuestionario

Los resultados del CVR figuran en la Table 41. De acuerdo con Lawshe (1975), el CVR aceptable para 15 expertos es de 0,49. Por consiguiente, los ítems con un CVR superior a 0,49 son coherentes y aceptables en términos de contenido con los objetivos del cuestionario. Los resultados muestran que se rechazaron 4 ítems, concretamente, PEU4, PEU7, PS3 y L5.

5.2.3 Índice de Validez de Contenido del Cuestionario

El Índice de Validez del Contenido (CVI) se calculó dos veces para el cuestionario: (1) pidiendo a las personas participantes que calificaran la pertinencia de cada ítem y (2) recabando su opinión sobre la claridad de cada uno de ellos.

5.2.3.1 Índice de validez del contenido del cuestionario a nivel de ítem (I-CVI) sobre la relevancia

Los resultados de la evaluación del CVI relacionados con la Relevancia se presentan en la Tabla 41. De acuerdo con las recomendaciones de Jahangiri et al. (2021), se fijó un valor mínimo aceptable de 0,79 para el I-CVI. Se determinó que los ítems con una puntuación I-CVI entre 0,7 y 0,79 requerían modificación, mientras que aquellos con una puntuación I-CVI inferior a 0,7 fueron rechazados. A partir del análisis, se rechazaron 7 ítems (PU8, PEU16, PS5, PS6, PS7, A4 y S4) y se identificaron 9 ítems (PU1, E2, L2, C9, A5, T2, S5, S10 y S13) que requerían modificación.

5.2.3.2 Índice de Validez del Contenido a Nivel de Ítem de Claridad (I-CVI) del Cuestionario

Los resultados de la evaluación del CVI relacionados con la Claridad se presentan en la Tabla 41. Basándose en las recomendaciones de Jahangiri et al. (2021), al igual que anteriormente, se fijó el valor mínimo aceptable de 0,79 para el I-CVI. Los ítems con una puntuación I-CVI entre 0,7 y 0,79 se identificaron como necesitados de modificación, mientras que los ítems con una puntuación I-CVI inferior a 0,7 se rechazaron. El análisis dio como resultado el rechazo de 10 ítems (PU1, PU9, PU11, PS5, C4, A4, A5, A6, T3 y S8) y la identificación de 10 ítems que requerían modificación (PU2, PU4, E1, C6, C7, A3, T2, S4, S6 y S7).

Tabla 41: Resultados de los cálculos de IIS, CVR y CVI de los ítems iniciales del cuestionario.

Constructo	Código	Ítem	Fase 3: Validez aparente		Fase 4: Ratio de Validez de Contenido		Fase 5: Índice de Validez de Contenido			
			IIS	Resultado	CVR	Resultado	Relevancia		Claridad	
							I-CVI	Resultado	I-CVI	Resultado
Utilidad percibida	PU1	El robot es útil.	3,36	Aceptado	1,00	Aceptado	0,73	Modificar	0,67	Rechazado
	PU2	El robot me ayuda a ser más eficaz.	3,31	Aceptado	0,87	Aceptado	0,93	Aceptado	0,73	Modificar
	PU3	El robot me ayuda a ser más productivo.	3,81	Aceptado	1,00	Aceptado	0,93	Aceptado	0,87	Aceptado
	PU4	El modo de control del robot mejora mi rendimiento laboral.	3,58	Aceptado	1,00	Aceptado	0,80	Aceptado	0,73	Modificar
	PU5	El robot me permite realizar rápidamente la tarea encomendada.	2,67	Aceptado	0,73	Aceptado	0,80	Aceptado	0,80	Aceptado
	PU6	El robot me ahorra tiempo cuando lo utilizo.	2,71	Aceptado	1,00	Aceptado	0,87	Aceptado	0,87	Aceptado
	PU7	El robot me da más control sobre las actividades de mi trabajo diario.	2,62	Aceptado	1,00	Aceptado	0,87	Aceptado	0,93	Aceptado
	PU8	El robot facilita las cosas que quiero hacer.	3,64	Aceptado	1,00	Aceptado	0,93	Aceptado	0,80	Aceptado
	PU9	Pude realizar movimientos precisos con el modo de control del robot.	2,32	Aceptado	0,87	Aceptado	0,67	Rechazado	0,67	Rechazado
	PU10	El robot satisface mis necesidades.	3,36	Aceptado	1,00	Aceptado	1,00	Aceptado	0,87	Aceptado
	PU11	El robot hace todo lo que esperaba que hiciera.	3,41	Aceptado	1,00	Aceptado	0,80	Aceptado	0,67	Rechazado
	PU12	Utilizar el robot mejora mi rendimiento en el trabajo.	4,36	Aceptado	1,00	Aceptado	1,00	Aceptado	0,93	Aceptado
Facilidad de uso percibida	PEU1	Mi interacción con el robot es sencilla.	4,36	Aceptado	1,00	Aceptado	0,93	Aceptado	0,80	Aceptado
	PEU2	El modo de control del robot es fácil de usar.	4,67	Aceptado	1,00	Aceptado	0,80	Aceptado	0,80	Aceptado
	PEU3	El modo de control del robot es sencillo de utilizar.	2,79	Aceptado	0,73	Aceptado	0,87	Aceptado	0,80	Aceptado
	PEU4	Utilizar el robot no supone ningún esfuerzo.	2,67	Aceptado	0,47	Rechazado	-	-	-	-
	PEU5	El modo de control del robot es fácil de usar.	3,70	Aceptado	0,73	Aceptado	0,80	Aceptado	0,87	Aceptado
	PEU6	*El modo de control del robot es rígido e inflexible.	1,46	Rechazado	-	-	-	-	-	-

Constructo	Código	Ítem	Fase 3: Validez aparente		Fase 4: Ratio de Validez de Contenido		Fase 5: Índice de Validez de Contenido			
			IIS	Resultado	CVR	Resultado	Relevancia		Claridad	
							I-CVI	Resultado	I-CVI	Resultado
	PEU7	*Este modo de control del robot resulta engorroso.	1,81	Aceptado	0,33	Rechazado	-	-	-	-
	PEU8	El modo de control del robot me resulta útil.	3,87	Aceptado	1,00	Aceptado	0,93	Aceptado	0,87	Aceptado
	PEU9	Siempre puedo utilizar el robot con éxito.	2,40	Aceptado	0,73	Aceptado	0,80	Aceptado	0,87	Aceptado
	PEU10	Puedo utilizar el robot sin instrucciones escritas.	3,76	Aceptado	1,00	Aceptado	0,87	Aceptado	1,00	Aceptado
	PEU11	Tanto a las personas usuarias ocasionales como a las habituales les gustaría que el robot.	1,46	Rechazado	-	-	-	-	-	-
	PEU12	No noto ninguna incoherencia al utilizar el robot.	2,12	Aceptado	0,73	Aceptado	0,80	Aceptado	0,80	Aceptado
	PEU13	El modo de control del robot requiere el menor número de pasos posible para lograr lo que quiero hacer con él.	3,64	Aceptado	1,00	Aceptado	0,80	Aceptado	0,93	Aceptado
	PEU14	El modo de control del robot permite recuperarse de los errores de forma rápida y sencilla.	4,04	Aceptado	1,00	Aceptado	0,93	Aceptado	0,93	Aceptado
	PEU15	Conseguir un objetivo concreto con la ayuda de un robot no será un problema para mí.	2,71	Aceptado	0,87	Aceptado	0,87	Aceptado	0,80	Aceptado
	PEU16	Podría hacer ajustes fáciles en un robot por mí mismo.	2,98	Aceptado	0,87	Aceptado	0,67	Rechazado	0,80	Aceptado
Seguridad percibida	PS1	Me siento seguro mientras utilizo el robot.	4,87	Aceptado	1,00	Aceptado	1,00	Aceptado	0,93	Aceptado
	PS2	Confiaba en que el robot no me haría daño.	4,17	Aceptado	0,87	Aceptado	0,87	Aceptado	1,00	Aceptado
	PS3	No me importa que el robot trabaje conmigo en un puesto de trabajo compartido.	1,81	Aceptado	0,47	Rechazado	-	-	-	-
	PS4	*Siento ansiedad al utilizar el robot.	2,93	Aceptado	0,87	Aceptado	0,80	Aceptado	0,80	Aceptado
	PS5	*Me siento agitado mientras utilizo el robot.	2,53	Aceptado	0,60	Aceptado	0,60	Rechazado	0,67	Rechazado
	PS6	El robot se movía demasiado rápido para mi comodidad.	2,88	Aceptado	0,73	Aceptado	0,60	Rechazado	0,93	Aceptado

Constructo	Código	Ítem	Fase 3: Validez aparente		Fase 4: Ratio de Validez de Contenido		Fase 5: Índice de Validez de Contenido			
			IIS	Resultado	CVR	Resultado	Relevancia		Claridad	
							I-CVI	Resultado	I-CVI	Resultado
	PS7	El robot se acercó demasiado a mí para mi comodidad.	2,88	Aceptado	0,73	Aceptado	0,60	Rechazado	0,93	Aceptado
Ergonomía	E1	Me siento físicamente incómodo al utilizar este modo de control.	3,70	Aceptado	0,73	Aceptado	0,87	Aceptado	0,73	Modificar
	E2	Me siento tenso al utilizar este modo de control.	3,52	Aceptado	0,73	Aceptado	0,73	Modificar	0,80	Aceptado
	E3	Me sentiría cómodo al interactuar con el robot.	4,17	Aceptado	1,00	Aceptado	0,93	Aceptado	0,80	Aceptado
Aprendibilidad	L1	Podría aprender fácilmente a utilizar un robot.	3,99	Aceptado	0,87	Aceptado	0,80	Aceptado	0,87	Aceptado
	L2	Aprendí a usarlo rápidamente.	4,04	Aceptado	0,87	Aceptado	0,73	Modificar	0,80	Aceptado
	L3	Recuerdo fácilmente cómo utilizarlo.	4,04	Aceptado	1,00	Aceptado	0,93	Aceptado	1,00	Aceptado
	L4	Rápidamente me volví hábil con él.	4,36	Aceptado	1,00	Aceptado	0,93	Aceptado	0,87	Aceptado
	L5	Puedo usar el robot si alguien me enseña cómo hacerlo primero.	3,98	Aceptado	0,47	Rechazado	-	-	-	-
Controlabilidad	C1	El robot es fácil de controlar.	4,42	Aceptado	1,00	Aceptado	1,00	Aceptado	0,93	Aceptado
	C2	Yo tenía el control general del robot.	4,42	Aceptado	1,00	Aceptado	1,00	Aceptado	1,00	Aceptado
	C3	Confío mucho en mi capacidad para controlar un robot.	3,87	Aceptado	1,00	Aceptado	1,00	Aceptado	1,00	Aceptado
	C4	Controlo el robot sobre el trabajo y el significado del trabajo.	2,24	Aceptado	0,73	Aceptado	0,80	Aceptado	0,60	Rechazado
	C5	Podría configurar un robot según mis deseos y mi entorno.	3,09	Aceptado	1,00	Aceptado	0,80	Aceptado	0,80	Aceptado
	C6	Creo que podría adaptar un robot para que me ayudara en mi vida diaria.	3,70	Aceptado	1,00	Aceptado	0,93	Aceptado	0,73	Modificar
	C7	Si tuviera que resolver un problema con la ayuda de un robot, podría hacerlo.	2,58	Aceptado	0,87	Aceptado	0,87	Aceptado	0,73	Modificar
	C8	Podría enseñarle algo a un robot si me esforzara lo suficiente.	1,49	Rechazado	-	-	-	-	-	-
	C9	Podría enseñar a un robot a realizar tareas sencillas.	2,58	Aceptado	1,00	Aceptado	0,73	Modificar	0,93	Aceptado
	C10	Si un robot hace algo mal, podría encontrar la manera de cambiar su	4,04	Aceptado	0,87	Aceptado	1,00	Aceptado	0,93	Aceptado

Constructo	Código	Ítem	Fase 3: Validez aparente		Fase 4: Ratio de Validez de Contenido		Fase 5: Índice de Validez de Contenido				
			IIS	Resultado	CVR	Resultado	Relevancia		Claridad		
							I-CVI	Resultado	I-CVI	Resultado	
		comportamiento.									
	C11	Podría desplegar un robot de una manera específica para ahorrar tiempo.	3,31	Aceptado	0,73	Aceptado	0,87	Aceptado	0,80	Aceptado	
Actitud e intención de uso	A1	Me vendría bien un robot en el trabajo diario.	3,47	Aceptado	1,00	Aceptado	1,00	Aceptado	0,93	Aceptado	
	A2	Podría hacer que un robot realizara una tarea específica.	3,87	Aceptado	1,00	Aceptado	0,93	Aceptado	0,93	Aceptado	
	A3	Si pudiera elegir que el robot me ayude en el trabajo, apreciaría trabajar con él.	3,81	Aceptado	1,00	Aceptado	0,93	Aceptado	0,73	Modificar	
	A4	Prefiero el robot a otras máquinas en el entorno industrial.	2,93	Aceptado	0,73	Aceptado	0,67	Rechazado	0,60	Rechazado	
	A5	El uso del robot es pertinente para mis diversas tareas laborales.	2,79	Aceptado	1,00	Aceptado	0,73	Modificar	0,60	Rechazado	
	A6	Creo que utilizar este modo de control es una buena idea.	1,88	Aceptado	1,00	Aceptado	0,80	Aceptado	0,67	Rechazado	
	A7	Creo que utilizaría este modo de control en futuras tareas.	2,99	Aceptado	0,73	Aceptado	0,87	Aceptado	0,80	Aceptado	
Confianza	T1	Confiaba en que el robot haría lo correcto en el momento adecuado.	4,04	Aceptado	1,00	Aceptado	0,93	Aceptado	0,87	Aceptado	
	T2	Si utilizara un robot, siempre sabría cómo y por qué se comporta como lo hace.	3,57	Aceptado	1,00	Aceptado	0,73	Modificar	0,73	Modificar	
	T3	Construí una relación significativa y útil con el robot.	1,96	Aceptado	0,87	Aceptado	0,80	Aceptado	0,47	Rechazado	
Satisfacción	S1	Estoy satisfecho con el robot.	4,23	Aceptado	1,00	Aceptado	0,93	Aceptado	0,93	Aceptado	
	S2	El robot y yo trabajamos bien juntos.	3,87	Aceptado	1,00	Aceptado	0,87	Aceptado	0,93	Aceptado	
	S3	Funciona como yo quiero.	4,36	Aceptado	0,87	Aceptado	0,87	Aceptado	0,87	Aceptado	
	S4	El robot se ajusta a mis expectativas personales.	2,49	Aceptado	1,00	Aceptado	0,67	Rechazado	0,73	Modificar	
	S5	Recomendaría el uso del robot a un colega.	4,23	Aceptado	0,87	Aceptado	0,73	Modificar	0,93	Aceptado	
	S6	El robot es agradable de usar.	3,52	Aceptado	1,00	Aceptado	0,80	Aceptado	0,73	Modificar	

Constructo	Código	Ítem	Fase 3: Validez aparente		Fase 4: Ratio de Validez de Contenido		Fase 5: Índice de Validez de Contenido			
			IIS	Resultado	CVR	Resultado	Relevancia		Claridad	
							I-CVI	Resultado	I-CVI	Resultado
	S7	*El robot no entendía cómo quería hacer la tarea.	2,76	Aceptado	0,87	Aceptado	0,87	Aceptado	0,73	Modificar
	S8	El robot seguía interponiéndose en mi camino.	3,31	Aceptado	0,87	Aceptado	0,80	Aceptado	0,67	Rechazado
	S9	*Los robots me hacen sentir incómodo.	3,36	Aceptado	0,87	Aceptado	0,87	Aceptado	0,87	Aceptado
	S10	Los robots me producen estrés adicional.	3,64	Aceptado	0,87	Aceptado	0,73	Modificar	0,80	Aceptado
	S11	Me gusta usar el robot.	3,08	Aceptado	1,00	Aceptado	0,80	Aceptado	0,93	Aceptado
	S12	Me siento cómodo con el robot.	3,64	Aceptado	0,87	Aceptado	0,87	Aceptado	0,87	Aceptado
	S13	*Siento ansiedad al utilizar el robot.	3,36	Aceptado	0,87	Aceptado	0,73	Modificar	0,93	Aceptado

5.3 Refinamiento de los ítems para evaluar la HRI en contextos industriales

Este apartado se enfoca en el refinamiento de los ítems que evalúan la utilidad percibida de un robot en el entorno laboral. A través de un proceso iterativo, se revisaron y modificaron los ítems originales para mejorar la claridad y precisión de la medición. El objetivo primordial radicó en optimizar la capacidad de la escala para capturar de manera efectiva la percepción de las personas usuarias en relación a la utilidad del robot en sus actividades laborales diarias.

5.3.1 Utilidad percibida

En relación a la dimensión de utilidad percibida, se contaba con un total de 9 ítems, de los cuales 7 ya habían sido aceptados y 2 requerían modificaciones. Con el fin de mejorar la claridad y garantizar una interpretación unívoca por parte de las personas participantes, se llevó a cabo una revisión exhaustiva de los ítems que conformaban la escala inicial de utilidad percibida.

Durante este proceso, se identificaron aquellos ítems que demandaban mayor clarificación y refinamiento, a fin de evitar ambigüedades en su interpretación. Notablemente, los ítems PU2, PU4 y PU12, orientados hacia un objetivo común de evaluar si el individuo percibe que el robot contribuye a su eficiencia, fueron fusionados en un único ítem. De manera análoga, los ítems PU5 y PU6 fueron amalgamados.

A continuación, se presenta la Tabla 42 que detalla las modificaciones implementadas en el constructo de Utilidad Percibida del cuestionario HUROY:

Tabla 42: Modificaciones realizadas en el constructo de Utilidad Percibida del cuestionario HUROY.

Código	Ítem
PU2 + PU4 + PU12	El robot mejora mi rendimiento laboral.
PU3	El robot me ayuda a ser más productivo/va.
PU5 + PU6	El robot me ayuda a completar la tarea en menos tiempo.
PU7	El robot me da más control sobre las actividades de mi trabajo diario.
PU8	El robot me ayuda a hacer las cosas más fácilmente.
PU10	El robot satisface mis necesidades.

5.3.2 Facilidad de uso percibida

En lo que concierne a la dimensión de la Facilidad de Uso Percibida, se había obtenido un conjunto de 11 ítems previamente aceptados. No obstante, con la finalidad de promover una optimización y una mayor concisión del cuestionario en sí, se propone la implementación de los ajustes detallados en la Tabla 43.

En particular, se ha fusionado el contenido de los ítems PEU1 y PEU2, en virtud de que ambos comparten el propósito de evaluar si el robot se presenta como una herramienta de uso sencillo. Asimismo, se ha efectuado una unión similar entre los ítems PEU9 y PEU12, puesto que ambas cuestionan la experiencia de uso satisfactorio del robot.

Tabla 43: Modificaciones realizadas en el constructo de Facilidad de Uso Percibida del cuestionario HUROY.

Código	Ítem
PEU1 + PEU 2	Utilizar el robot es fácil.
PEU3	El modo de control del robot es sencillo de utilizar.

PEU5	El modo de control del robot es amigable.
PEU8	El modo de control del robot me resulta útil.
PEU9 + PEU12	Siempre puedo utilizar el robot con éxito.
PEU10	Puedo utilizar el robot sin instrucciones escritas.
PEU13	El modo de control del robot requiere el menor número de pasos posible para lograr lo que quiero hacer con él.
PEU14	El modo de control del robot permite recuperarse de los errores de forma rápida y sencilla.
PEU15	Conseguir un objetivo concreto con la ayuda de un robot no será un problema para mí.

5.3.3 Seguridad percibida

En el marco del constructo de Seguridad Percibida, se ha depurado el conjunto de ítems a tres, como resultado de la evaluación experta (Tabla 44). Se considera que estos tres ítems poseen la suficiente claridad y pertinencia para evaluar de manera adecuada la dimensión de seguridad percibida.

Por otro lado, durante el análisis del constructo "Ergonomía", intrínsecamente vinculado a la seguridad percibida, se ha constatado la necesidad de introducir modificaciones sustanciales. Conforme a esta constatación, se ha tomado la determinación de suprimir el constructo "Ergonomía" en su formato previo, y en su lugar, incorporar el ítem pertinente dentro del constructo de Seguridad Percibida. Además, se ha optado por una reformulación del enunciado de dicho ítem en términos negativos. Este enfoque, que contrasta con la formulación positiva original ("He experimentado confort al utilizar el robot"), se ajusta de manera más precisa a la realidad industrial, donde la prioridad radica en la minimización de la percepción de incomodidad o malestar durante la interacción con el robot.

Tabla 44: Modificaciones realizadas en el constructo de Seguridad percibida y Ergonomía del cuestionario HUOX

Código	Ítem
PS1	Me siento seguro/ra mientras utilizo el robot.
PS2	Confiaba en que el robot no me haría daño.
PS4	*Sentí ansiedad mientras utilizaba el robot.
C1	*Experimenté incomodidad al utilizar el robot.

5.3.4 Controlabilidad

Dentro del ámbito de la evaluación de la controlabilidad, se contemplaron un total de ocho ítems, de los cuales seis habían sido previamente validados por las personas expertas. Sin embargo, dos de estos ítems demandaron ajustes orientados a mejorar su nivel de claridad y comprensión. Con este objetivo, se procedió a la revisión minuciosa y a la optimización de la redacción de dichos ítems, con el fin de favorecer su interpretación por parte de las personas usuarias. Los cambios implementados se encuentran detallados en la Tabla 45.

Tabla 45: Modificaciones realizadas en el constructo de Controlabilidad del cuestionario HUOX.

Código	Ítem
C1	El robot es fácil de controlar.
C2	Siento que tengo el control del robot.
C3	Tengo mucha confianza en mi capacidad para controlar el robot.
C5	Puedo configurar el robot según mis necesidades.
C6	Puedo ajustar el robot para que me ayude a realizar mis tareas.
C7	Puedo resolver problemas con la ayuda del robot.

C10	Si el robot hace algo mal, puedo encontrar la manera de cambiar su comportamiento.
C11	Puedo desplegar el robot de una manera específica para ahorrar tiempo.

5.3.5 Aprendibilidad

Dentro del constructo de la Aprendibilidad, si bien los cuatro ítems habían recibido la aprobación de las personas expertas, se determinó que el ítem L5 adolecía de falta de claridad. Esta carencia implicaba que, dicho ítem, en su formulación actual, no aportaba un contenido de valor substancial al cuestionario. La afirmación "Podría usar el robot fácilmente si alguien me enseña primero cómo hacerlo" podía conllevar ambigüedades interpretativas. Por lo tanto, en aras de garantizar la integridad de la medición y la coherencia en la evaluación de la Aprendibilidad, se optó por excluir este ítem.

En consecuencia, conforme se observa en la Tabla 46, el constructo de Aprendibilidad se encuentra ahora representado de manera efectiva por tres ítems que han demostrado poseer una redacción clara y comprensible.

Tabla 46: Modificaciones realizadas en el constructo de Aprendibilidad del cuestionario HUROX.

Código	Ítem
L1	Puedo aprender fácilmente a utilizar el robot.
L3	Recuerdo fácilmente cómo utilizar el robot.
L4	Rápidamente adquirí destreza con el robot.

5.3.6 Actitud

En cuanto a la Actitud e Intención de Uso hacia el sistema, la evaluación experta reveló que, de los cuatro ítems inicialmente propuestos, uno de ellos requería modificaciones con el propósito de mejorar su claridad. En tal contexto, y como evidencia la Tabla 47, se procedió a la optimización del ítem A3 a fin de propiciar una comprensión más efectiva por parte de las personas participantes.

Además, enriqueciendo este constructo, se ha incorporado un nuevo ítem que originalmente formaba parte del constructo Confianza. Tras la evaluación experta, se determinó que este ítem, en su esencia, evalúa la confianza en el sistema, la cual es una manifestación de la actitud hacia el sistema. Esta adición se consideró esencial para capturar más plenamente la intersección entre la actitud y la confianza en el contexto de la interacción con el robot.

Tabla 47: Modificaciones realizadas en el constructo de Actitud del cuestionario HUROX.

Código	Ítem
A1	Podría utilizar el robot en mi trabajo.
A2	Podría hacer que el robot realizara una tarea específica.
A3	Si tuviera la opción, preferiría trabajar con el robot
A7	Creo que utilizaría este modo de control en una tarea futura
T1	Confío en que el robot haga lo correcto en el momento adecuado.

5.3.7 Satisfacción

Dentro del constructo de Satisfacción, de ocho ítems seis contaban con la aceptación de las personas expertas, mientras que dos requerían ajustes. Específicamente, los ítems S6 y S7 fueron objeto de modificación en aras de mejorar su coherencia y relevancia.

Como reflejo de este proceso, la Tabla 48 ilustra la modificación del ítem S6 para optimizar su comprensión y su alineación con la dimensión de Satisfacción. Asimismo, se ha

tomado la decisión de eliminar el ítem S7, debido a su limitado valor añadido y escasa contribución a la evaluación integral de la satisfacción.

Tabla 48: Modificaciones realizadas en el constructo de Satisfacción del cuestionario HUROX.

Código	Ítem
S1	Estoy satisfecho/cha con el robot.
S2	El robot y yo hemos trabajado bien juntos/as.
S3	El robot funciona como yo quiero que funcione.
S6	El robot es agradable de usar.
S9	El robot me hace sentir incómodo/da.
S11	Me gusta utilizar el robot.
S12	Me siento cómodo/da con el robot.

5.4 Discusión y conclusiones

El cuestionario HUROX, que se adjunta en el Anexo B, incorpora en su estructura un conjunto de 7 constructos, abarcando un total de 41 ítems. Estos constructos son, la utilidad percibida (6 ítems), facilidad de uso percibida (9 ítems), seguridad percibida (4 ítems), controlabilidad (8 ítems), aprendibilidad (3 ítems), actitud hacia el sistema (5 ítems) y satisfacción (7 ítems). Estos, se miden en una escala de Likert de 7 puntos.

La literatura actual carece de cuestionarios que aborden de manera integral la percepción de las personas en entornos de interacción humano-robot (HRI) industriales. Con el propósito de cerrar esta brecha, se ha propuesto el cuestionario HUROX como una herramienta destinada a mejorar la evaluación de la experiencia del usuario (UX), la aceptación de la tecnología y el bienestar general mediante la recopilación de datos perceptuales. HUROX aborda esta limitación al incorporar los factores más influyentes en la adopción de robots, conforme al marco conceptual HRIX.

Este cuestionario desempeña un papel fundamental en la implementación del modelo ITPX-ROBOTS, ya que se alinea de manera coherente con la tercera fase de este modelo, que se centra en la recopilación de datos perceptuales para una comprensión integral de la HRI.

La versatilidad y adaptabilidad de HUROX permiten su aplicación en una amplia gama de entornos industriales, lo que lo convierte en una herramienta prometedora para la evaluación de la percepción de las personas usuarias en la interacción con robots. Sin embargo, se requiere un proceso de validación adicional y la recopilación de datos empíricos en múltiples contextos para garantizar su fiabilidad y validez. Además, la retroalimentación de las personas usuarias finales será esencial para refinar aún más este cuestionario y asegurar su eficacia en la evaluación de la experiencia en la HRI.

Principales conclusiones del capítulo

- Los cuestionarios de UX son herramientas valiosas en el campo de la HRI en entornos industriales. A pesar de su potencial, es evidente la falta de cuestionarios específicos para evaluar la UX y la aceptación de la tecnología en contextos de HRI industriales. **La creación y validación de un cuestionario como HUROY** se convierte en un paso esencial hacia la mejora del diseño y la aceptación de sistemas HRI en este contexto. El cuestionario completo se muestra en el Anexo A.
- Combinar cuestionarios con otras técnicas de evaluación objetiva o cualitativa, como se propone en el modelo ITPX-ROBOTS, brinda una visión más integral de cómo las personas que utilizan el robot en estos contextos la perciben y la aceptan. Esto permite obtener una comprensión más completa de la UX y la aceptación de la tecnología en estos entornos.
- **El cuestionario HUROY se compone de 7 categorías principales**, cada una de las cuales incluye un conjunto de preguntas específicas para un total de 41 preguntas en total. Estas categorías se enfocan en evaluar la **utilidad percibida** (6 preguntas), la **facilidad de uso percibida** (9 preguntas), la **seguridad percibida** (4 preguntas), la **controlabilidad** (8 preguntas), la **aprendibilidad** (3 preguntas), la **actitud hacia el sistema** (5 preguntas) y la **satisfacción** (7 preguntas). Las respuestas a estas preguntas se registran en una escala de 7 puntos de Likert.

Capítulo VI

Validación del modelo ITPX-ROBOTS

6 VALIDACIÓN DEL MODELO ITPX-ROBOTS

En este capítulo, se presentan los resultados de las dos experimentaciones diseñadas con el propósito de validar las hipótesis planteadas en el estudio (Tabla 49). En congruencia con el objetivo central de esta tesis, *"Diseñar, definir y validar un nuevo modelo de evaluación centrado en UX para optimizar la interacción entre personas y robots en entornos industriales"*, estos casos de estudio se motivan principalmente por la necesidad de poner a prueba y validar el modelo ITPX-ROBOTS. El propósito primordial radica en determinar si dicho modelo es efectivo y pertinente en contextos prácticos.

El enfoque principal de estos casos de estudio reside en evaluar la capacidad del modelo ITPX-ROBOTS para capturar y comprender la dinámica de la interacción entre personas y robots en entornos industriales. Además, se persigue la confirmación de la Hipótesis 1 planteada, la cual establece que *"En el contexto de la HRI en entornos industriales, un aumento en el indicador HEUROBOX se traducirá en un incremento en los niveles de valencia emocional y engagement, mientras que los niveles de memorización y carga mental se verán reducidos."* La Hipótesis 1 será validada mediante la exploración de la correlación del indicador HEUROBOX con la valencia emocional, el *engagement*, la memorización y la carga mental en ambos casos de estudio.

Por lo tanto, estos casos de estudio se erigen como un pilar fundamental para asegurar que el modelo ITPX-ROBOTS sea un instrumento eficaz y confiable en la evaluación de la UX y adopción de robots en entornos de HRI en la industria, lo que potencialmente contribuirá a una mejor comprensión y optimización de estas interacciones.

Tabla 49: Descripción de los casos de estudio.

N.º del caso de estudio	Objetivo	Robot utilizado	Indicadores de desempeño	Dispositivo fisiológico utilizado	Cuestionario	Muestra	Hipótesis que validar
1	Evaluar la respuesta en las variables cinemáticas	KUKA LBR iiwa 14 R820	Tiempo Errores	EEG y GSR	UEQ-S	20	Hipótesis 1 Hipótesis 2
2	Evaluar la respuesta en interfaces unimodales o multimodales	KUKA LBR iiwa 14 R820	Índice de producción	EEG	HUROX	20	Hipótesis 1 Hipótesis 3

La Hipótesis 2 plantea que *"En el contexto de la interacción persona-robot en un entorno industrial, cuanto más elevados sean los valores de velocidad y aceleración del robot, incrementará la carga mental, mientras que los valores de la valencia emocional y la percepción general de la experiencia de uso serán disminuidos."* Para abordar esta hipótesis, se ha diseñado específicamente el primer caso de estudio, que, además de validar el modelo ITPX-ROBOTS, se enfoca en evaluar las respuestas emocionales y cognitivas de las personas usuarias ante las diferentes variables cinemáticas mientras colaboran con un robot en una tarea de ensamblado. Los detalles de la metodología y los resultados de este primer caso de estudio se presentan en la sección 6.2.

La Hipótesis 3 plantea que *"En el contexto de interacción persona-robot en un entorno industrial, el uso de interfaces multimodales conlleva una disminución en los valores de memorización y carga mental y un incremento de la experiencia de uso percibida en comparación con el uso de una interfaz unimodal."* Para abordar esta hipótesis, se ha diseñado específicamente el segundo caso de estudio, que, además de validar el modelo ITPX-

ROBOTS, se enfoca en evaluar las respuestas emocionales y cognitivas de las personas usuarias ante diferentes interfaces (unimodales vs. multimodales) durante la colaboración con un robot en una tarea de desensamblado. Los detalles de la metodología y los resultados de este segundo caso de estudio se presentan en la sección 6.3.

6.1 Entorno y equipamiento utilizado

A continuación, se describen el entorno y el equipamiento utilizados para llevar a cabo los dos casos de estudio abordados.

6.1.1 Entorno de los casos de estudio

Los dos casos de estudio se llevaron a cabo en un entorno de laboratorio que proporcionó un ambiente controlado y propicio para la investigación. Este enfoque permitió una concentración precisa en las variables bajo estudio, que incluyen las variables cinemáticas en la percepción de las personas en el primer caso de estudio, y la influencia de las interfaces unimodales o multimodales en el segundo.

El laboratorio utilizado se encuentra en la Escuela Politécnica Superior de Mondragon Unibertsitatea y se conoce como el Laboratorio de Industria 4.0. Este espacio de investigación abarca una superficie de 112 m² y cuenta con equipamiento tecnológico de última generación como visión artificial, un Robot Colaborativo, 2 Robots industriales, un sistema modular de producción y un robot móvil. Toda esta infraestructura de sistemas industriales se encuentra a su vez conectada a un Centro de Proceso de Datos (Mondragon Unibertsitatea, 2020).

6.1.2 Robot utilizado: KUKA LBR iiwa 14 R820

En el laboratorio de industria 4.0 se encuentra el robot KUKA IIWA (Figura 40). Se trata de un robot sensible producido en serie y, por lo tanto, compatible con el HRC. Con el LBR iiwa, las personas y los robots pueden trabajar juntos en tareas altamente sensibles en estrecha colaboración, lo que abre la posibilidad de nuevas aplicaciones y allana el camino hacia una mayor rentabilidad y eficiencia (*LBR iiwa | KUKA AG, s. f.*).



Figura 40: Robot utilizado en los casos de estudio: KUKA IIWA 14 R820 (*LBR iiwa | KUKA AG, s. f.*).

6.1.3 Dispositivos fisiológicos

A continuación, se describen de forma detallada los dispositivos fisiológicos utilizados en los casos de estudio, cada uno diseñado para capturar aspectos específicos de las respuestas emocionales y cognitivas de las personas participantes.

Para empezar, Diadem es un EEG (Bitbrain, 2019) de sensores secos diseñado para aplicaciones del mundo real que requieren, por un lado, gran comodidad para la persona usuaria, y por otro, una configuración ágil y una calidad de señal sobresaliente para el

investigador. Diadem ha sido desarrollado con 12 sensores secos ubicados en áreas cerebrales específicas (pre-frontal, frontal, parietal y occipital) para la estimación de estados emocionales y cognitivos. Su apantallamiento activo de alto rendimiento y su diseño mecánico estable proporcionan una robustez y una calidad de señal excepcionales, incluso bajo movimiento o durante largos periodos de grabación (Figura 41).



Figura 41: Diadem de Bitbrain (2019).

Por otro lado, Ring (Figura 42) de Bitbrain (2019) es un dispositivo de medición de bioseñales móvil y robusto para la monitorización de la conductancia de la piel (sensor electrodérmico – EDA/medidor respuesta galvánica de la piel – GSR) y la actividad cardiovascular mediante un sensor de pulso de volumen sanguíneo (BVP, por sus siglas en inglés).

Ha sido diseñado para su uso en contextos reales de investigación que requieren una gran comodidad para la persona usuaria, una configuración ágil y una calidad de señal sobresaliente para el investigador. El dispositivo tiene un diseño ultraligero y cómodo, y dispone de dos biosensores clave (actividad electrodérmica mediante GSR y actividad cardíaca mediante BVP) para la estimación de estados emocionales, y un acelerómetro solidario de tres ejes, para una estimación del ruido que se puede generar debido al movimiento de la mano.



Figura 42: Ring de Bitbrain (2019).

6.2 Caso de estudio 1: Influencia de las variables cinemáticas

El propósito de este caso de estudio es examinar el impacto de las variables cinemáticas, en particular la velocidad y la aceleración, en los aspectos relacionados con los factores humanos y la percepción de las personas usuarias. Para lograrlo, se emplea el modelo ITPX-ROBOTS. En las siguientes secciones, se describirá en detalle el procedimiento seguido y se presentarán los resultados obtenidos en cada etapa.

6.2.1 Motivación

Las variables cinemáticas de un robot, en particular la velocidad y la aceleración, desempeñan un papel importante en la percepción que las personas tienen de dicho robot (Arai et al., 2010; Dehais et al., 2011; Dragan et al., 2015; Hanajima et al., 2017; Lasota & Shah, 2015). Sin embargo, es relevante destacar que la consideración de la velocidad y la aceleración de los movimientos ha estado principalmente orientada hacia la seguridad en el contexto de los robots industriales. Esta orientación ha llevado a que los aspectos emocionales de las personas sean pasados por alto en el proceso de definir las variables cinemáticas del robot (Koppenborg et al., 2017).

Koppenborg et al. (2017) estudiaron los efectos de la velocidad de movimiento y la previsibilidad de la trayectoria en las personas operarias, utilizando un entorno de RV. En sus hallazgos, revelaron que los movimientos más rápidos estaban asociados con evaluaciones más elevadas de carga de trabajo mental y niveles de ansiedad. En los estudios de Lasota & Shah (2015) y Beschi et al. (2020) también abordaron las implicaciones del movimiento del robot, aunque no examinaron de manera exhaustiva los aspectos cognitivos y perceptuales involucrados. Rubagotti et al. (2021) identificaron factores clave relacionados con el movimiento y las características del robot que influyen en la percepción de seguridad, concluyendo, entre otras cosas, que la percepción de seguridad aumenta a medida que disminuye la velocidad del robot. De acuerdo con los hallazgos de Joosse et al. (2021), una posible razón subyacente a la preferencia de las personas por una velocidad más lenta en los robots radica en la falta de claridad en cuanto a cuándo el robot tiene la intención de detenerse, es decir, la ausencia de *feedback* que comunique su intencionalidad. Esto sugiere la oportunidad de investigar las variables cinemáticas y su correlación con la percepción y el desempeño de las personas usuarias.

La realización de este caso de estudio se encuentra motivada por la necesidad de profundizar en la comprensión del impacto de las variables cinemáticas, en particular la velocidad y la aceleración, en aspectos cruciales como la carga mental, la valencia emocional y la percepción general de la experiencia de las personas usuarias en interacciones con robots.

Este estudio busca arrojar luz sobre cómo las características cinemáticas de los robots pueden influir significativamente en la experiencia y percepción de las personas usuarias, así como en factores humanos clave. La comprensión de esta influencia desempeñará un papel fundamental en la toma de decisiones de las personas desarrolladoras e investigadoras, permitiéndoles ajustar las variables cinemáticas y diseñar robots de acuerdo con las preferencias individuales y las necesidades de las personas usuarias.

6.2.2 Metodología

La metodología empleada, en consonancia con el modelo ITPX-ROBOTS, consiste en una evaluación experta mediante la herramienta HEUROBOX, y a continuación un test de usuario que implica la recopilación de indicadores de desempeño e indicadores fisiológicos durante la interacción con el robot, así como la recopilación de datos perceptuales posterior a la interacción.

6.2.2.1 Evaluación experta a través de HEUROBOX

La evaluación previa al test de usuarios basada en heurísticos corresponde a la primera fase del modelo ITPX-ROBOTS. De acuerdo con el modelo ITPX-ROBOTS, se ha procedido a llevar a cabo una rigurosa evaluación del sistema robótico bajo análisis, empleando la herramienta de evaluación heurística experta denominada HEUROBOX.

6.2.2.2 Procedimiento experimental en el Test de Usuario

Este caso de estudio tiene como objetivo esclarecer la influencia de las variables cinemáticas, específicamente la velocidad y la aceleración, en los factores humanos y en la percepción de las personas usuarias. Para ello, se utiliza el modelo ITPX-ROBOTS. De acuerdo con el modelo y las directrices planteadas por Apraiz, Lasa, Montagna, et al. (2023), el procedimiento experimental se ha estructurado en tres fases distintas, como se muestra en la Figura 43: (i) la fase previa a la ejecución de la tarea, (ii) la fase de ejecución de la tarea y (iii) la fase posterior a la ejecución de la tarea.

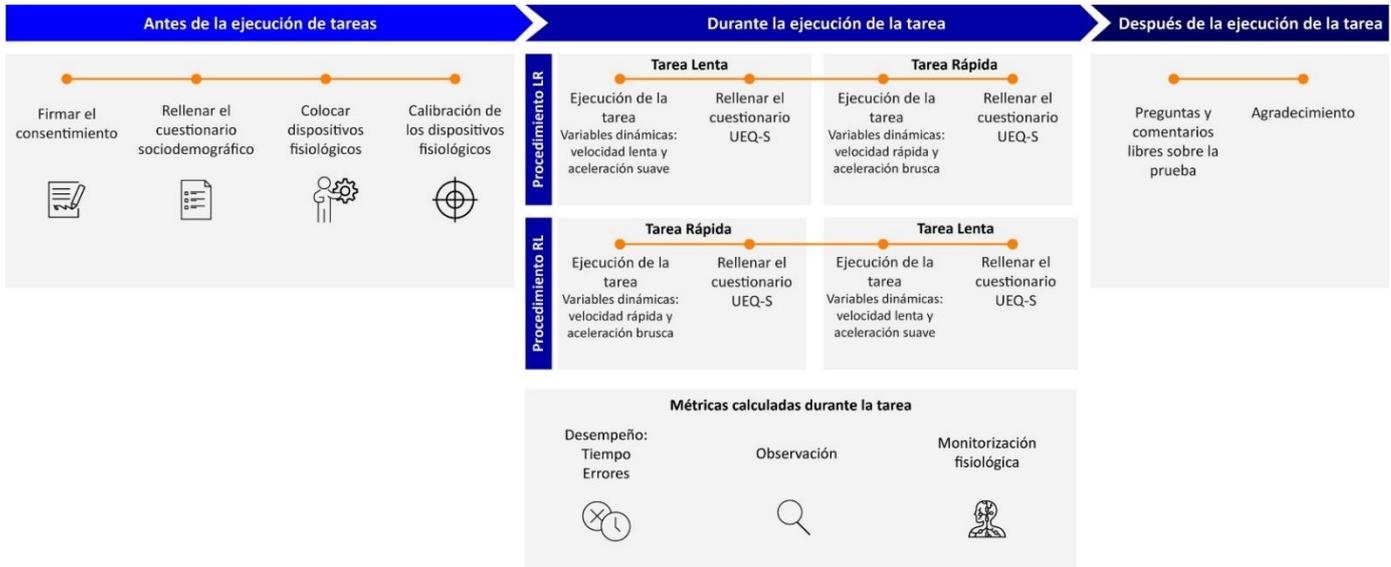


Figura 43: Procedimiento experimental seguido en el caso de estudio 1.

6.2.2.2.1 Fase previa a la ejecución de la tarea

En esta etapa inicial (Figura 44), se procedió a obtener el consentimiento informado de las personas participantes, detallando tanto la naturaleza como el alcance de la prueba. Esto incluyó la recopilación de datos, que abarcó la grabación de imágenes y sonidos, así como la monitorización biométrica mediante EEG y GSR. Posteriormente, las personas participantes completaron un cuestionario sociodemográfico diseñado para recopilar información relevante, como la edad y el género. Además, se proporcionaron instrucciones detalladas sobre cómo interactuar con el robot a través de comandos de voz. Una vez completadas estas etapas, se procedió a colocar a las personas participantes los dispositivos fisiológicos, específicamente el Diadem y el Ring de Bitbrain (cuyos detalles se explicarán posteriormente), y se procedió a su calibración siguiendo el protocolo validado de calibración del *software* SennsLab (Bitbrain Technologies, 2018).



Figura 44: Procedimiento experimental ANTES de la ejecución de tareas en el caso de estudio 1.

6.2.2.2.2 Durante la fase de ejecución de la tarea

Una vez calibrados los dispositivos fisiológicos, se diferenciaron dos procedimientos (Figura 45):

- **En el procedimiento Lento-Rápido (LR)**, las personas participantes (1-10) completaron primero la tarea en un modo de movimiento lento del robot (Tarea Lenta, TL), es decir, la velocidad del robot era lenta (62,5 mm/s) y la aceleración suave (25%). Una vez finalizada la tarea, rellenaron el cuestionario UEQ-S. A continuación, repitieron la misma tarea en modo de movimiento rápido (Tarea Rápida, TR), es decir, con la velocidad del robot alta (187,5 mm/s) y la aceleración alta (75%). A continuación, rellenaron de nuevo el cuestionario UEQ-S.
- **En el procedimiento Rápido-Lento (RL)**, las personas participantes (11-20) realizaron primero la tarea en el modo de movimiento rápido (TR). A continuación, rellenaron el cuestionario UEQ-S. Después de rellenar el cuestionario, realizaron la tarea en el modo de movimiento lento (TL). A continuación, rellenaron el cuestionario UEQ-S.

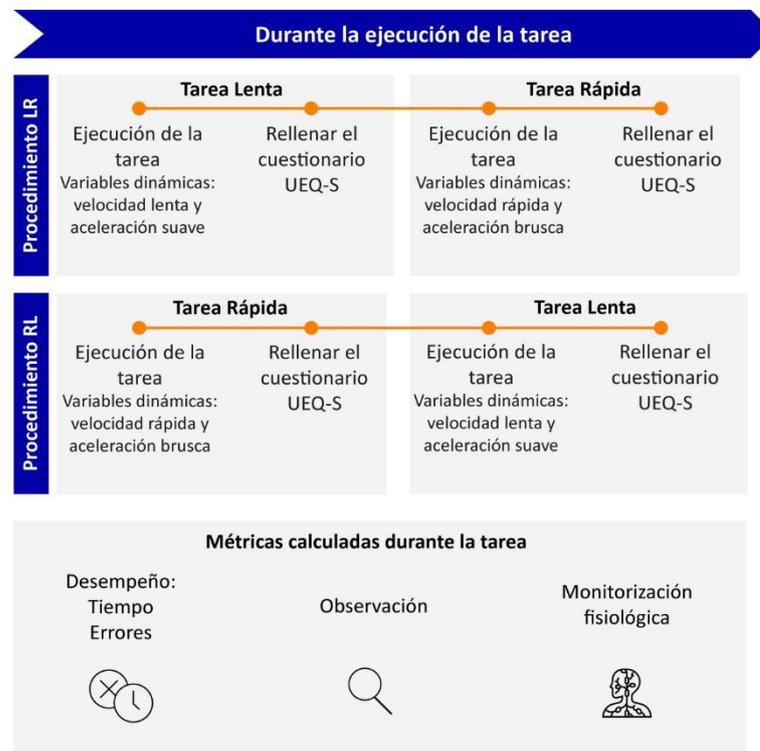


Figura 45: Procedimiento experimental DURANTE la ejecución de las tareas en el caso de estudio 1.

6.2.2.2.3 Fase posterior a la ejecución de la tarea

Después de la interacción, se llevó a cabo una conversación libre con las personas participantes (Figura 46). Se les formularon preguntas acerca de sus emociones y expectativas, con el propósito de fomentar un ambiente en el que se sintieran cómodos expresando sus sentimientos y pensamientos libremente.

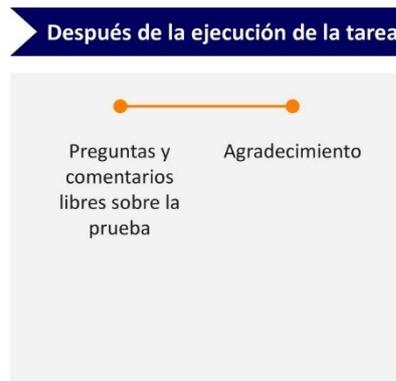


Figura 46: Procedimiento experimental DESPUÉS de la ejecución de la tarea en el caso de estudio 1.

6.2.2.3 Tarea

Se desarrolló una tarea de ensamblaje colaborativo que requería la participación activa de las personas participantes. En esta tarea, las personas participantes fueron encomendadas con la selección y posicionamiento de tres componentes: un cilindro, un cubo y una base, que debían ensamblar. Un factor crucial consistía en asegurar que la base fuera correctamente colocada sobre un marco. Dado que la base presentaba dos orificios con formas geométricas distintas, uno para cada componente, se requería la precisa ubicación del cilindro y el cubo en sus respectivas posiciones (Figura 47). Posteriormente, a través de una interfaz de voz, las personas participantes instruían al robot para que presionara las dos piezas hacia abajo, logrando así su ensamblaje con la base. Una vez completado el ensamblaje, el participante procedía a retirar la base, ahora unida a las dos piezas, del marco.

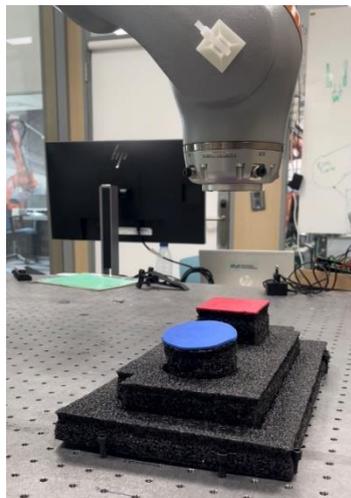


Figura 47: Tarea del robot y contexto.

Este proceso se repetía hasta en cinco ocasiones. Para brindar una visualización más detallada de la secuencia de actividades durante la tarea, se presenta la Figura 48.

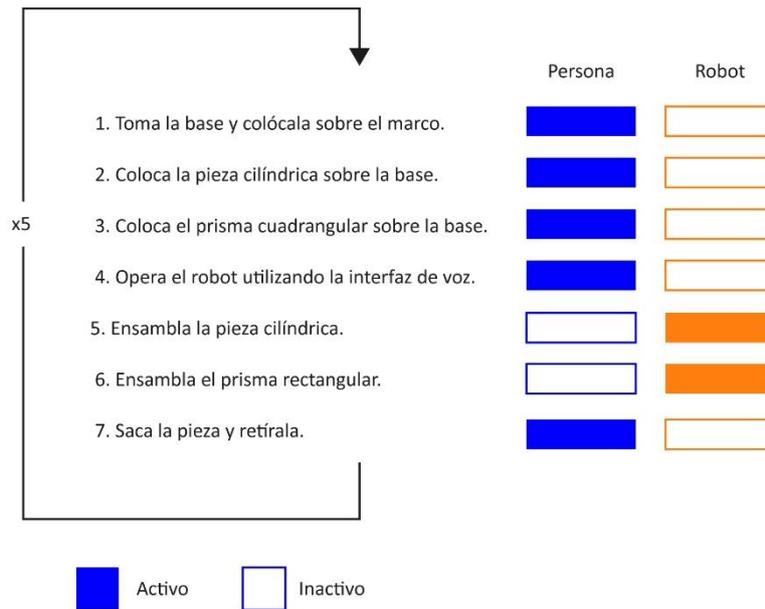


Figura 48: Secuencia de tareas en un ensamblaje del caso de estudio 1.

6.2.2.4 Indicadores del experimento

Al implementar el modelo ITPX-ROBOTS, los indicadores que se han utilizado son: (i) indicadores de desempeño (tiempo de ejecución de la tarea y errores), que son medidas cuantitativas y objetivas; (ii) valores fisiológicos, que son cuantitativos y objetivos; y (iii) valores perceptivos, que son cuantitativos y subjetivos.

6.2.2.4.1 Indicadores de desempeño

Dada la naturaleza y el planteamiento de la tarea, se ha optado por registrar dos indicadores clave de desempeño: (i) el tiempo de ejecución de la tarea, que abarca el periodo desde la asignación de la tarea hasta su finalización, y (ii) los errores, todas aquellas acciones que no estaban contempladas en la secuencia de acciones previamente definida. Es importante destacar que, en este caso particular, el cálculo del índice de producción no resulta pertinente, ya que a las personas participantes se les solicitó realizar un total de 5 ensamblados.

6.2.2.4.2 Indicadores fisiológicos

Los indicadores fisiológicos se han recogido mediante los dispositivos Diadem y Ring de Bitbrain (detallados en el apartado 6.1.3). El Diadem proporciona los siguientes indicadores:

- i) **Valencia (EEG):** Grado de atracción experimentado durante la presencia de un estímulo o una situación. Varía desde un extremo “positivo/agradable” hasta un extremo “negativo/desagradable” (Bitbrain, 2020).
- ii) **Memorización (EEG):** Este indicador mide la intensidad de los procesos cognitivos relacionados con la formación de futuros recuerdos durante la presentación de un estímulo. Captura el grado de almacenamiento, codificación y retención en la memoria (Bitbrain, 2020).
- iii) **Atención (EEG):** Este indicador mide la focalización (hablando siempre en términos de procesos neuronales) de un participante cuando se le presenta un estímulo. Es decir, la utilización de los recursos cognitivos para la realización de una tarea o la visualización de un estímulo (Bitbrain, 2020).

- iv) **Engagement (EEG):** Es un indicador del grado de implicación o conexión entre el participante y el estímulo o tarea. No es simple atención ya que el participante puede estar totalmente atento y que la información que se presenta no sea de su interés y por lo tanto no exista dicha conexión o involucración (Bitbrain, 2020).

Por su parte, el Ring proporciona las siguientes métricas:

- i) **Activación (Ring):** Nivel basal de activación fisiológica producido por un estímulo o situación. Valores menores que 0 se asocian a un estado de relajación o calma y valores por encima de 0 se asocian a un estado de excitación. Notar que la activación emocional puede ser debida a una emoción positiva o a una emoción negativa. La activación se expresa en porcentajes (Bitbrain, 2020).
- ii) **Impacto (Ring):** El impacto emocional mide el número y la intensidad de los cambios puntuales del estado emocional producidos por una estimulación, evento externo o la realización de una tarea (algo que llama especialmente la atención, produce excitación o estrés (Bitbrain, 2020).

6.2.2.4.3 Indicadores perceptuales

Los valores de percepción de las personas usuarias se recogieron mediante el cuestionario UEQ-S (Schrepp et al., 2017). El cuestionario contiene dos subescalas: calidad pragmática (CP) y calidad hedónica (CH). También refleja el valor global de la UX. Además, en el contexto de la HRI, la seguridad percibida, la confianza y la utilidad percibida son factores humanos relevantes (Coronado et al., 2022; Hancock et al., 2011; Onnasch & Roesler, 2021). Como el cuestionario UEQ-S no los incluye, se han añadido tres ítems para evaluarlos también en una escala Likert de 7 puntos:

- **Fiabilidad:** Poco fiable - Fiable.
- **Controlabilidad:** Incontrolable - Controlable.
- **Utilidad percibida:** No valiosa - Valiosa.

Los conceptos de fiabilidad y controlabilidad son medidas importantes para evaluar la percepción del robot por parte de la persona usuaria. La fiabilidad se refiere a la consistencia y fiabilidad del robot, mientras que la controlabilidad se refiere a la facilidad de manejo y la capacidad de la persona usuaria para controlar el robot. Estos factores desempeñan un papel crucial a la hora de determinar el nivel de confianza y seguridad percibida en la interacción. Unas puntuaciones altas en Fiabilidad y Controlabilidad indican que las personas usuarias confían en que el robot funcione de forma coherente y pueden controlarlo fácilmente, lo que genera una sensación de seguridad. Por el contrario, puntuaciones bajas en estos conceptos pueden provocar una falta de confianza y una menor percepción de seguridad, lo que puede repercutir en la experiencia global. Por otro lado, la utilidad percibida se refiere al grado en que un individuo cree que el uso del robot mejorará su rendimiento laboral o aumentará su productividad. Refleja el grado en que el participante considera que el robot es útil y añade valor a su experiencia. Es un factor crucial para determinar la satisfacción general del participante y su disposición a seguir utilizando el robot.

6.2.2.5 Procesamiento de los datos

En la Tabla 50 se resumen las variables de respuesta e independientes consideradas en el estudio. Los datos fisiológicos se recogieron con Sennslab, un *software* de registro de datos multimodal (Bitbrain, 2019). A continuación, se procesaron los datos mediante Sennsmetrics (Bitbrain, 2019) una extensión de Sennslab que permite analizar los datos fisiológicos.

Tabla 50: Resumen de las variables del caso de estudio 1.

Tipo de variable	Categoría de la variable	Variable	Descripción
Variable independiente	Factores de configuración	Variables cinemáticas	Velocidad y aceleración (baja, alta).
Variables de respuesta	Indicadores de desempeño	Tiempo de ejecución de la tarea	Desde el momento en que se asigna la tarea hasta su finalización.
		Errores	Todas aquellas acciones que no estaban previstas en la secuencia.
	Indicadores perceptuales	Calidad Pragmática (CP)	Percepción de la capacidad de apoyar la consecución de los "do-goals" (Hassenzahl, 2007).
		Calidad Hedónica (CH)	Percepción de la capacidad de apoyar la consecución de los "be-goals" (Hassenzahl, 2007).
		Fiabilidad	Confianza percibida hacia el robot.
		Controlabilidad	Percepción de la capacidad de controlar el robot.
		Utilidad percibida	El grado de convencimiento de que su uso mejorará su rendimiento o aumentará su productividad.
	Respuestas fisiológicas	Valencia	El grado de atracción experimentado (Bitbrain, 2021).
		Memorización	La intensidad de los procesos cognitivos relacionados con la formación de recuerdos futuros (Bitbrain, 2021).
		Carga mental	El uso de recursos cognitivos para llevar a cabo una tarea (Bitbrain, 2021).
Engagement		El grado de implicación o conexión entre el participante y la tarea (Bitbrain, 2021).	
Activación	Nivel basal de Activación fisiológica (Bitbrain, 2021).		
Impacto	El número y la intensidad de los cambios puntuales del estado emocional (Bitbrain, 2021).		

6.2.2.6 Muestra

En el experimento participaron 20 personas voluntarias (10 hombres y 10 mujeres) distribuidas equitativamente en ambos procedimientos:

- En el procedimiento LR participaron diez personas de entre 21 y 41 años (media = 28,7, std=6,7). Cuatro de ellos tenían experiencia previa con un robot doméstico, cinco con robots industriales y uno nunca había interactuado con un robot.
- En el procedimiento de RL participaron diez personas de edades comprendidas entre los 24 y los 50 años (media = 30,5 std=7,8). Dos participantes habían interactuado previamente con un robot doméstico, cuatro con un robot industrial y cuatro nunca habían interactuado con un robot.

La mano dominante de todas las personas participantes era la derecha, por lo que todos llevaban el anillo en la izquierda.

6.2.3 Resultados

La estructuración del análisis de datos sigue la metodología de investigación propuesta de acuerdo con el modelo ITPX-ROBOTS, dividiendo los resultados en cuatro secciones fundamentales:

1. Indicador HEUROBOX
2. Indicadores de desempeño.
3. Indicadores fisiológicos.
4. Indicadores perceptuales.

6.2.3.1 Indicador HEUROBOX

Esta evaluación se ha llevado a cabo utilizando la versión avanzada de HEUROBOX. Como resultado de esta evaluación, se ha obtenido un indicador promedio del 47,46%.

Al profundizar en el análisis de los resultados obtenidos, se observa el siguiente desglose (Tabla 51). En particular, la seguridad y la ergonomía han demostrado ser áreas sólidas con valores del 52% y 69,56% respectivamente, indicando una adecuada atención a la protección y comodidad de la persona usuaria. Sin embargo, la funcionalidad presenta un nivel de 27,27%, señalando posibles áreas de mejora en la eficacia y capacidad operativa del sistema. Las interfaces también requieren atención, evidenciado por un puntaje del 40%, sugiriendo la necesidad de optimización en la interacción entre la persona usuaria y el sistema. El análisis global del indicador arroja un promedio ponderado del 47,46%, subrayando la importancia de un enfoque holístico para la mejora continua y la optimización del sistema robótico evaluado.

Tabla 51: Resultado del indicador HEUROBOX en el caso de estudio 1.

Categoría	Resultado (%)
Seguridad	52%
Ergonomía	69,56%
Funcionalidad	27,27%
Interfaces	40%
TOTAL INDICADOR	47,46%

6.2.3.2 Indicadores de desempeño

Los indicadores de desempeño se analizaron midiendo el tiempo de ejecución de la tarea y los errores. En la TL, el robot operó a una velocidad de 62,5 mm/s, y el tiempo total requerido para ejecutar la TL sin la intervención humana fue de 2 minutos y 28 segundos. En la TR, el robot funcionó a una velocidad de 187,5 mm/s, y el tiempo total necesario para completar la TR sin tener en cuenta el tiempo de intervención humana fue de 1 minuto y 20 segundos.

Los resultados de los indicadores de desempeño se presentan en la Tabla 52 para el procedimiento LR y en la Tabla 54 para el procedimiento RL. Analizando los valores absolutos sin tener en cuenta el orden de ejecución de las tareas, los resultados muestran que las personas participantes tardaron menos tiempo en completar la TR. Estos resultados son lógicos porque, como el robot actuaba más rápido, las personas participantes tenían más posibilidades de terminar antes la tarea.

Si se analizan los resultados teniendo en cuenta el orden de las tareas, en el procedimiento LR, todas las personas participantes tardaron más tiempo en completar la TL que las personas participantes en el procedimiento RL. Del mismo modo, el tiempo para completar la TR fue menor para las personas participantes del procedimiento LR que para los del procedimiento RL. Por lo tanto, se observa que las personas participantes que realizaron las tareas en el orden inverso de los procedimientos (LR y RL) tardaron más tiempo en

completar las tareas y cometieron más errores en la primera tarea que en la segunda. Esto indica que el orden en que se realizan las tareas puede influir en los resultados y sugiere que las personas participantes pueden tener que familiarizarse más con la tarea para conseguir un mejor rendimiento.

Si se analiza en función del género, en términos de desempeño no se observan diferencias significativas en el procedimiento LR. Las mujeres que participaron en el procedimiento LR necesitaron en promedio 3:23 minutos para completar la TL, mientras que los hombres requirieron en promedio 3:14 minutos, una diferencia que no es estadísticamente significativa. En la TR, las mujeres emplearon en promedio 2:07 minutos, mientras que los hombres utilizaron en promedio 2:05 minutos, nuevamente sin diferencias significativas en el tiempo de ejecución basadas en el género. En cuanto a los errores, las mujeres promediaron 1,4 errores en la TL, en contraste con los hombres que promediaron 2,2 errores. En la TR, tanto hombres como mujeres cometieron un promedio de 0,2 errores, sin diferencias significativas basadas en el género.

Tabla 52: Indicadores de desempeño en el procedimiento LR.

Id	Género	TL		TR	
		Tiempo	Errores	Tiempo	Errores
1	Hombre	3:16	1	2:10	0
2	Mujer	3:22	2	2:01	0
3	Mujer	3:12	2	2:03	1
4	Hombre	3:26	1	2:19	0
5	Hombre	2:49	2	1:43	1
6	Mujer	3:28	2	2:06	0
7	Mujer	3:37	0	2:15	0
8	Hombre	3:20	2	2:11	0
9	Hombre	3:23	5	2:06	0
10	Mujer	3:18	1	2:12	0

Las diferencias entre los resultados obtenidos en ambas tareas son significativas de acuerdo con los resultados del T-test (Tabla 53), tanto en el tiempo de ejecución de la tarea como en la cantidad de errores cometidos.

Tabla 53: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR.

Tiempo	Errores
6,65E-11	3,85E-03
Diferencia significativa	Diferencia significativa

En relación con los resultados obtenidos en el procedimiento RL, tampoco se observan diferencias significativas en función del género. Las mujeres que participaron en el procedimiento RL necesitaron en promedio 2:26 minutos para completar la TR, mientras que los hombres requirieron en promedio 2:10 minutos, una diferencia que no alcanza la significación estadística. En la TL, las mujeres emplearon en promedio 3:08 minutos, mientras que los hombres utilizaron en promedio 3:00 minutos, nuevamente sin diferencias significativas en el tiempo de ejecución basadas en el género. En cuanto a los errores, las mujeres promediaron 1 error en la TR, en contraste con los hombres que promediaron 2,6 errores. En la TL, las mujeres hicieron un promedio de 0,4 errores, mientras que los hombres cometieron 1,4 errores en promedio.

Tabla 54: Indicadores de desempeño en el procedimiento RL.

Id	Género	TR		TL	
		Tiempo	Errores	Tiempo	Errores
11	Hombre	2:14	1	3:13	1
12	Hombre	2:17	4	3:00	3
13	Mujer	2:30	1	3:00	0
14	Mujer	2:33	3	3:15	2
15	Mujer	2:27	0	3:09	0
16	Mujer	2:23	1	3:19	0
17	Mujer	2:17	0	3:01	0
18	Hombre	2:00	3	2:44	2
19	Hombre	2:00	3	2:56	1
20	Hombre	2:23	2	3:09	0

Como se evidencia en la Tabla 55, se observan diferencias significativas entre los resultados obtenidos en ambas tareas, tanto en el tiempo de ejecución de la tarea como en la cantidad de errores cometidos.

Tabla 55: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento RL.

Tiempo	Errores
2,20E-08	1,40E-01
Diferencia significativa	Diferencia significativa

En promedio, las personas participantes en el procedimiento LR tomaron 48 segundos más de lo esperado para completar la TL, y las personas participantes en el procedimiento RL también requirieron, en promedio, 36 segundos más de lo esperado para ejecutar TL. En lo que respecta a la TR, las personas participantes en el procedimiento LR tomaron, en promedio, 45 segundos más de lo previsto, mientras que los del procedimiento RL necesitaron, en promedio, 58 segundos adicionales.

En cuanto a los errores, el orden en el que se realizaron las tareas tuvo un impacto significativo en el rendimiento. En el procedimiento LR, las personas participantes cometieron más errores en TL que en TR, mientras que en el procedimiento RL, ocurrió lo contrario: cometieron más errores en TR que en TL.

Por lo tanto, se concluye que el orden de realización de las tareas afecta los resultados durante su ejecución. La segunda tarea se completó en ambos procedimientos con un mejor desempeño que la primera, y la cantidad de errores también disminuyó en la tarea realizada en segundo lugar. Sin embargo, es importante destacar que el tiempo necesario para completar la tarea fue más rápido en TR. En este sentido, el factor de orden tiene un impacto significativo en la capacidad de aprendizaje, ya que puede influir considerablemente en la velocidad y eficacia con la que las personas usuarias son capaces de aprender y utilizar eficazmente el sistema.

6.2.3.3 Indicadores fisiológicos

Esta sección muestra y analiza los resultados obtenidos con los dispositivos fisiológicos. A continuación, se muestran y analizan los resultados obtenidos mediante el EEG en el Procedimiento LR (Figura 49):

- **Engagement.** Se observó un nivel de *engagement* mayor en las personas participantes en la TR (M=47,81; SD=57,40) en comparación con la TL

(M=33,81; SD=4,53). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 56).

- **Memorización:** Se observó que las personas participantes tuvieron levemente una menor memorización en la TL (M=30,40; SD=6,00) que en la TR (M=31,03; SD=5,05). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 56).
- **Valencia:** La valencia emocional se registró como más negativa en la TL (M=-3,04; SD=7,62) en contraste con la TR (M=-1,89; SD=9,76). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 56).
- **Carga mental:** Se observó que la carga mental fue menor en la TL (M=29,58; SD=3,10) en comparación con la TR (M=37,21; SD=19,59). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 56).
- **Activación:** Se observa que la activación ha sido mayor en la TL (M=18,75; SD=16,15) que en la TR (M=-0,74; SD=27,59). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 56).
- **Impacto:** Se observa que el impacto ha sido menor en la TL (M=34,94; SD=1,68) que en la TR (M=36,43; SD=2,55). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 56).

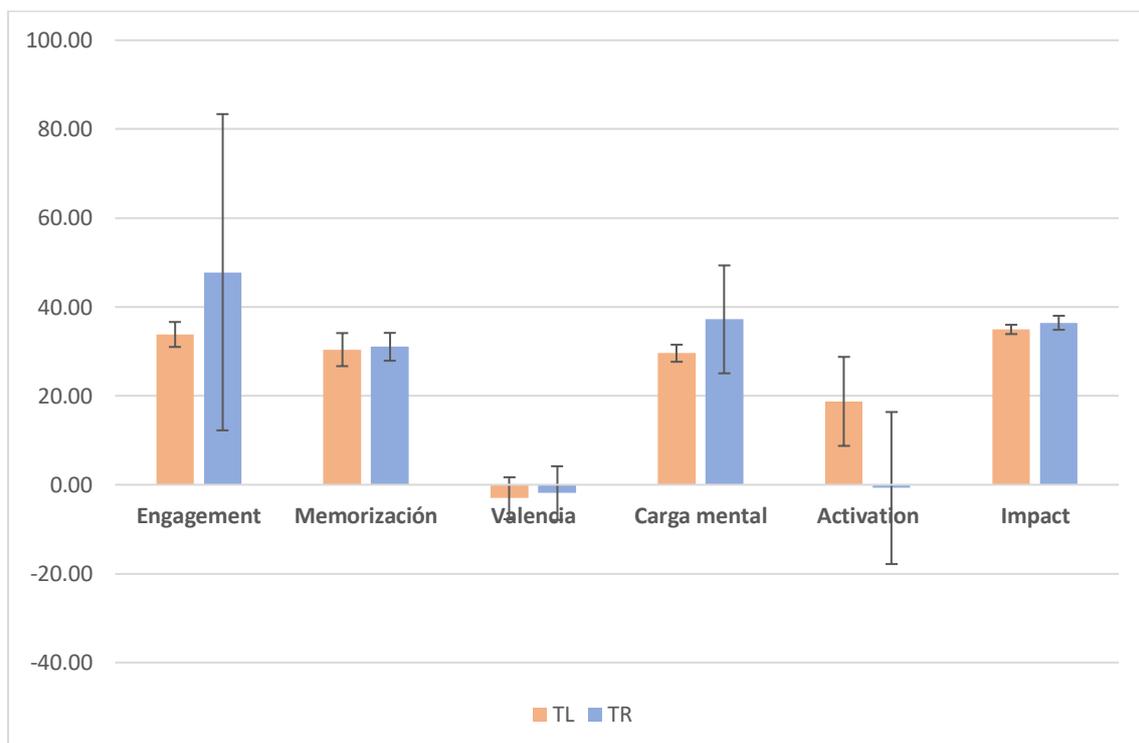


Figura 49: Resultados de los indicadores fisiológicos en el procedimiento LR en el caso de estudio 1.

Sin embargo, las diferencias mostradas en la Figura 48, no son estadísticamente significativas (Tabla 56).

Tabla 56: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR.

<i>Engagement</i>	Memorización	Valencia	Carga mental	Activación	Impacto
0,4613	0,8011	0,7722	0,2532	0,0737	0,1420
No hay diferencia significativa					

Desagregando por género los datos del procedimiento LR, obtenemos las siguientes resultados (Figura 50):

- **Engagement:** En el caso de las mujeres se observa que el nivel de *engagement* ha sido ligeramente mayor en la TL (M=32,45; SD=5,50) que en la TR (M=29,44; SD=5,72). Sin embargo, en el caso de los hombres se observa que el nivel de *engagement* ha sido significativamente menor en la TL (M=35,16; SD=3,36) que en la TR (M=66,18; SD=80,85). Sin embargo, las diferencias no son estadísticamente significativas (Tabla 57).
- **Memorización:** En el caso de las mujeres se observa que la memorización ha sido ligeramente menor en la TL (M=30,27; SD=7,28) que en la TR (M=31,96; SD=6,06). Al contrario, en el caso de los hombres se observa que la memorización ha sido ligeramente superior en la TL (M=30,53; SD=5,43) que en la TR (M=30,11; SD=4,31). Sin embargo, las diferencias no son estadísticamente significativas (Tabla 57).
- **Valencia:** En el caso de las mujeres se observa que la valencia emocional es mejor en la TL (M=-1,07; SD=9,84) que en la TR (M=-3,31; SD=2,79). Sin embargo, en el caso de los hombres la valencia emocional es inferior en la TL (M=-5,01; SD=4,29) que en la TR (M=-0,47; SD=14,19). Sin embargo, las diferencias no son estadísticamente significativas (Tabla 57).
- **Carga mental:** En la TL, se observa que la carga mental es inferior tanto en hombres como en mujeres. Sin embargo, la diferencia es más pronunciada en los hombres, donde la carga mental es mayor en la TR en comparación con la TL. Los datos muestran que, en mujeres, en la TL (M=28,77; SD=2,07) y en la TR (M=29,89; SD=3,95), mientras que, en hombres, en la TL (M=30,39; SD=3,96) y en la TR (M=44,52; SD=26,72). Sin embargo, las diferencias no son estadísticamente significativas (Tabla 57).
- **Activación:** Tanto en hombres como en mujeres se observa una mayor activación en la TL (Mujeres: M=24,32; SD=18,88; Hombres: M= 13,17; SD=12,37) que en la TR (Mujeres: M= 16,63; SD=18,62; Hombres: M= -18,11; SD=24,74). En mujeres, esta diferencia no es estadísticamente significativa, pero en hombres sí se considera estadísticamente significativa (Tabla 57).
- **Impacto:** Tanto en hombres como en mujeres se observa un menos impacto en la TL (Mujeres: M=35,14; SD=1,93; Hombres: M= 34,74; SD=1,58) que en la TR (Mujeres: M= 35,77; SD=1,16; Hombres: M= 37,09; SD=3,49). Sin embargo, las diferencias no son estadísticamente significativas (Tabla 57).

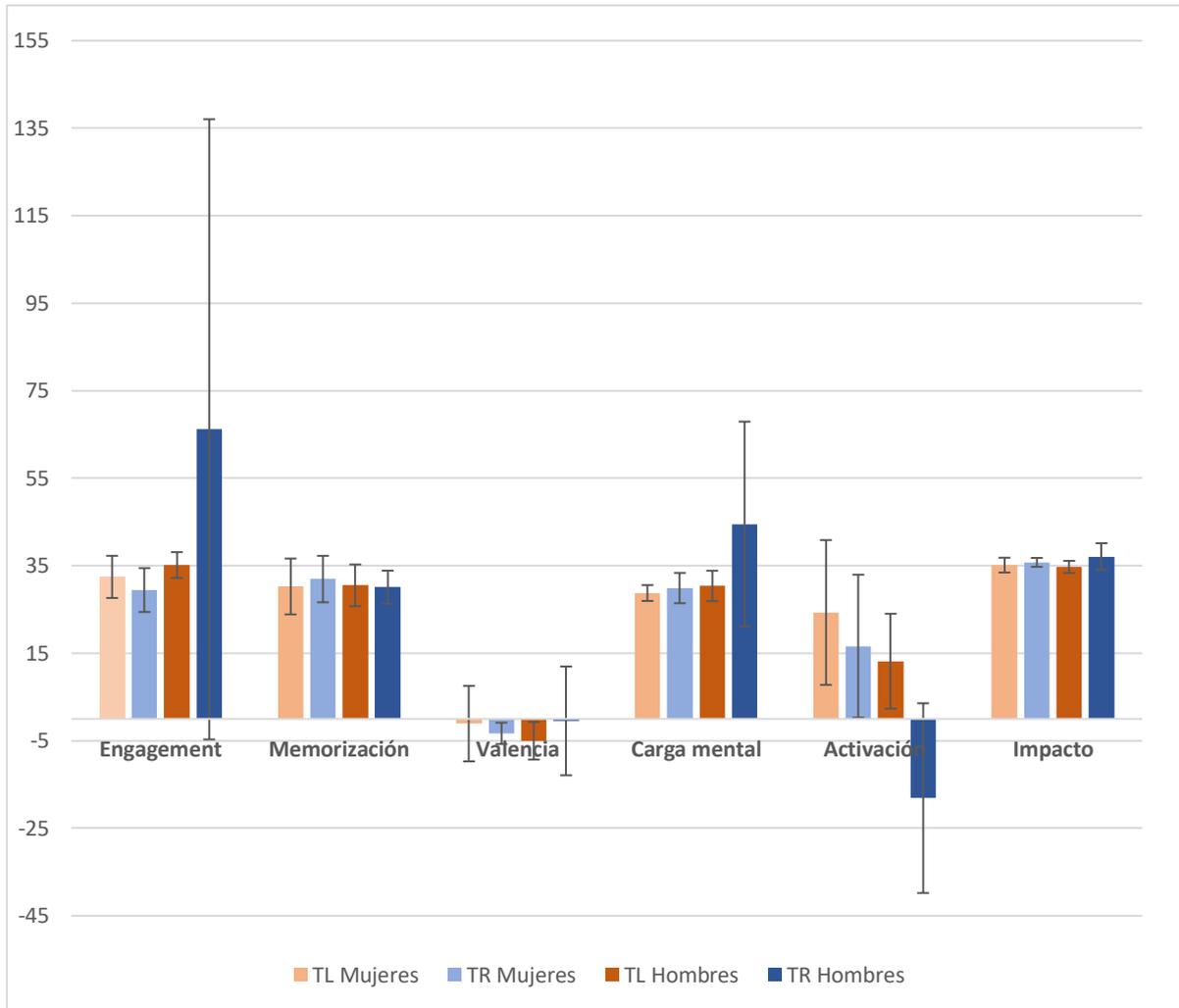


Figura 50: Resultados de los indicadores fisiológicos en el procedimiento LR en el caso de estudio 1 desagregados por género.

Por lo tanto, tal y como se observa en la Tabla 57, la única diferencia significativa es la que corresponde a los hombres en los indicadores de la activación, ya que han obtenido valores menores de activación en la TR.

Tabla 57: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR desagregado por género.

	Engagement	Memorización	Valencia	Carga mental	Activación	Impacto
Mujeres	0,4214 No hay diferencia significativa	0,6981 No hay diferencia significativa	0,6474 No hay diferencia significativa	0,5934 No hay diferencia significativa	0,5347 No hay diferencia significativa	0,5540 No hay diferencia significativa
Hombres	0,4396 No hay diferencia significativa	0,8953 No hay diferencia significativa	0,5294 No hay diferencia significativa	0,3043 No hay diferencia significativa	0,0455 Diferencia significativa	0,2218 No hay diferencia significativa

A continuación, se muestran y analizan los resultados obtenidos mediante el EEG y el GSR en el Procedimiento RL (Figura 51):

- **Engagement.** Se observó un nivel de *engagement* menor en las personas participantes en la TR (M=32,84; SD=4,67) en comparación con la TL (M=36,04;

SD=7,71). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 58)

- **Memorización:** Se observó que las personas participantes tuvieron mayor nivel de memorización en la TR (M=32,12; SD=4,69) que en la TL (M=23,74; SD=4,93). Esta diferencia se considera estadísticamente significativa (Tabla 58).
- **Valencia:** La valencia emocional se registró como más negativa en la TR (M=-3,08; SD=6,19) en contraste con la TL (M=5,00; SD=19,63). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 58).
- **Carga mental:** Se observó que la carga mental fue similar en la TR (M=30,35; SD=3,68) en comparación con la tarea TL (M=30,53; SD=3,97). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 58).
- **Activación:** Se observa que la activación ha sido menos negativa en la TL (M=-0,71; SD=34,29) que en la TR (M=-9,79; SD=31,84). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 58).
- **Impacto:** Se observa que el impacto ha sido menor en la TR (M=35,06; SD=1,71) que en la TR (M=35,50; SD=4,17). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 58).

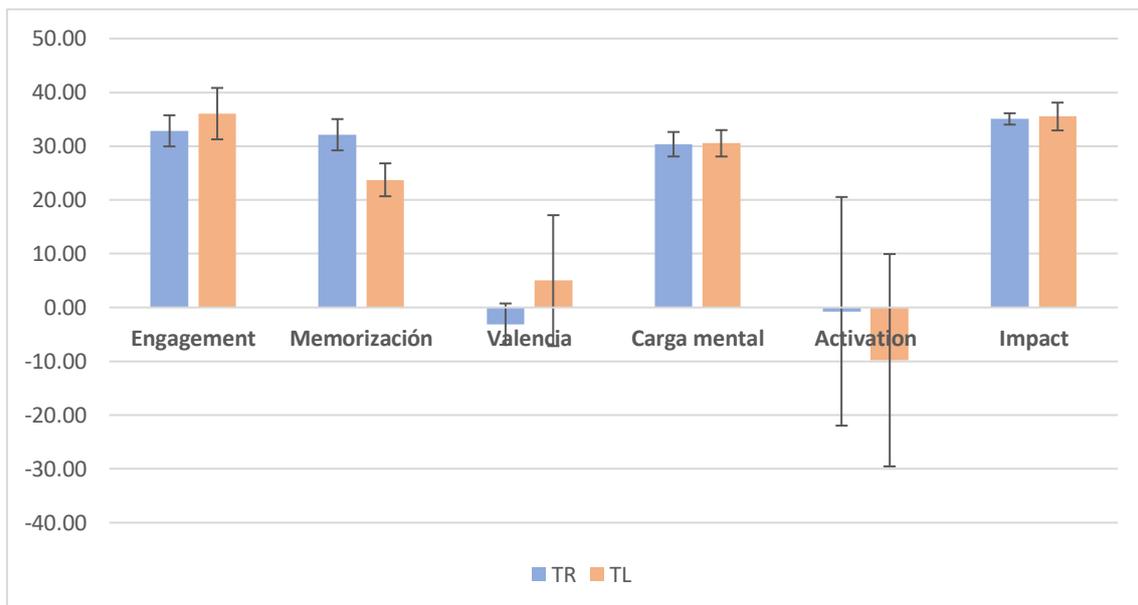


Figura 51: Resultados de los indicadores fisiológicos en el procedimiento RL en el caso de estudio 1.

Sin embargo, es importante destacar que, según los resultados de la Tabla 58, solo la diferencia en los indicadores de memorización resulta significativa entre las tareas.

Tabla 58: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento RL.

Engagement	Memorización	Valencia	Carga mental	Activación	Impacto
0,2801	0,0011	0,2408	0,9208	0,5468	0,7587
No hay diferencia significativa	Diferencia significativa	No hay diferencia significativa			

Desagregando por género los datos del procedimiento RL, se han obtenido los siguientes resultados (Figura 52):

- **Engagement:** Tanto en hombres como en mujeres se observa un menor nivel de *engagement* en la TR (Mujeres: $M=33,75$; $SD=4,17$; Hombres: $M= 31,93$; $SD=5,45$) que en la TL (Mujeres: $M=35,04$; $SD=9,54$; Hombres: $M= 37,04$; $SD=6,35$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 59).
- **Memorización:** Tanto en hombres como en mujeres se observa un mayor nivel de memorización en la TR (Mujeres: $M=30,35$; $SD=3,40$; Hombres: $M= 33,89$; $SD=5,49$) que en la TL (Mujeres: $25,24$; $SD=5,69$; Hombres: $M= 22,24$; $SD=4,10$). En mujeres, la diferencia no se considera estadísticamente significativa, pero en hombres sí (Tabla 59).
- **Valencia:** Tanto en hombres como en mujeres se observa una menor valencia emocional en la TR (Mujeres: $M=-2,90$; $SD=4,15$; Hombres: $M=-3,27$; $SD=8,29$) que en la TL (Mujeres: $M=7,92$; $SD=28,64$; Hombres: $M= 2,07$; $SD=5,03$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 59).
- **Carga mental:** En el caso de las mujeres, se observa una mayor carga mental en la TR ($M=29,61$; $SD=3,45$) que en la TL ($M=28,67$; $SD=2,29$). Sin embargo, en el caso de los hombres se observa una menor carga mental en la TR ($M=31,10$; $SD=4,14$) que en la TL ($M=32,38$; $SD=4,64$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 59).
- **Activación:** En el caso de las mujeres se observa una activación menos negativa en la TR ($M=-13,05$; $SD=32,48$) que en la TL ($M= -23,09$; $SD=42,27$). En el caso de los hombres, se observa una mayor activación en la TR ($M=11,64$; $SD=34,78$) que en la TL ($M=3,50$; $SD=7,21$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 59).
- **Impacto:** En el caso de las mujeres se observa un impacto similar tanto en la TR ($M=34,56$; $SD=1,88$) que en la TL ($M=34,38$; $SD=1,68$). En el caso de los hombres se observa un menor impacto en la TR ($M=35,56$; $SD=5,75$) que en la TL ($M=-36,63$; $SD=5,75$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 59).

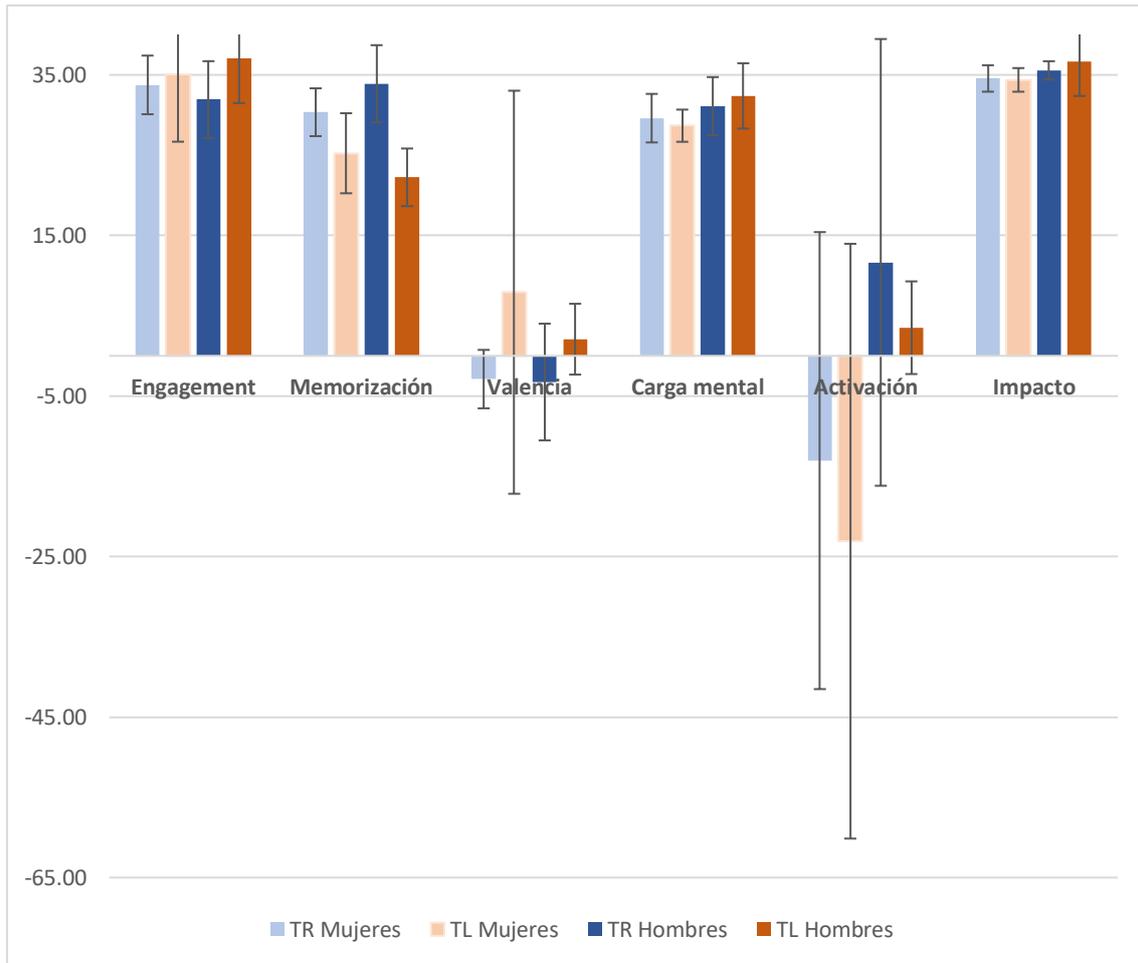


Figura 52: Resultados de los indicadores fisiológicos en el procedimiento RL en el caso de estudio 1 desagregados por género.

Sin embargo, es relevante destacar que, de acuerdo con los resultados de la Tabla 59, solo se observa una diferencia significativa entre las tareas en los indicadores de memorización en los hombres.

Tabla 59: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento RL desagregado por género.

	<i>Engagement</i>	Memorización	Valencia	Carga mental	Activación	Impacto
Mujeres	0,7929	0,1313	0,4485	0,6275	0,6855	0,8774
	No hay diferencia significativa					
Hombres	0,2100	0,0060	0,2605	0,6568	0,6334	0,7046
	No hay diferencia significativa	Diferencia significativa	No hay diferencia significativa			

6.2.3.3.1 Conclusiones generales de los resultados en los indicadores fisiológicos

En cuanto a las conclusiones generales de este estudio, se pueden extraer varios hallazgos:

- **Influencia de las variables cinemáticas en la valencia emocional:** Los resultados sugieren que la velocidad y la aceleración del robot tienen un impacto ligeramente negativo en la valencia emocional de las personas participantes. Específicamente, se observa que las personas participantes experimentaron emociones más negativas cuando el robot se movía a velocidades más altas. Esto indica que las variables cinemáticas, como la velocidad y la aceleración, pueden influir en el estado emocional durante la HRI.
- **Impacto de las variables cinemáticas en la carga mental:** Se identifica la influencia de la velocidad y la aceleración del robot en la carga mental de las personas participantes. En general, las personas participantes experimentan una mayor carga de trabajo cuando el robot se mueve a velocidades más altas. Además, se observó que los hombres tendían a experimentar una carga de trabajo mayor que las mujeres, independientemente de la velocidad y la aceleración del robot. Estos resultados resaltan la importancia de ajustar las variables cinemáticas para optimizar la carga de trabajo de las personas usuarias en entornos industriales.
- **Relación entre variables cinemáticas y *engagement*:** Se han encontrado diferencias en el *engagement* de las personas participantes en función de las variables cinemáticas. En general, las personas participantes muestran niveles más altos de *engagement* cuando el robot opera a velocidades y aceleraciones más bajas. Esto sugiere que las variables cinemáticas influyen en el *engagement*, y que una velocidad más lenta y una aceleración suave promueven un mayor *engagement* en la HRI.
- **Impacto de la duración de la interacción en la activación y el impacto:** Se han observado diferencias en las puntuaciones medias de activación e impacto de las personas participantes en relación con la duración de la interacción con el robot. En el procedimiento LR, tanto mujeres como hombres han experimentado niveles más altos de activación en la TL que en la TR. Por otro lado, en el procedimiento RL, las personas participantes han experimentado niveles más altos de activación e impacto en la TR que en la TL. Estos resultados sugieren que la duración de la interacción con el robot puede influir en las respuestas de activación e impacto de las personas usuarias.

6.2.3.4 Indicadores perceptuales

En esta sección, se lleva a cabo un análisis descriptivo de los indicadores perceptuales. Para una visión general de las diferencias entre las distintas tareas en cada aspecto evaluado en el cuestionario, se presentan las Figuras 52 y 54, correspondientes al procedimiento LR y RL, respectivamente. Además, para obtener una comprensión más detallada, los valores concretos se detallan en la Tabla 64.

La Figura 53 muestra los resultados obtenidos en el procedimiento LR. Además, tal y como puede observarse en la Tabla 54, los resultados de la CP se identifican como una diferencia significativa.

- **CP:** Se observó un nivel de CP menor en las personas participantes en la TL ($M=5,45$; $SD=0,71$) en comparación con la TR ($M=6,075$; $SD=0,54$). Esta diferencia se considera estadísticamente significativa (Tabla 60).

- **CH:** Se observó un nivel de CH menor en las personas participantes en la TL (M=4,5; SD=1,09) en comparación con la TR (M=4,925; SD=1,66). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 60).
- **Fiabilidad:** Se observó un nivel mayor de fiabilidad en la TL (M=6,4; SD=1,26) en comparación con la TR (M=5,8; SD=1,55). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 60).
- **Controlabilidad:** Se observó un nivel mayor de controlabilidad en la TL (M=5,5; SD=2,01) en comparación con la TR (M=4,8; SD=1,93). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 60).
- **Utilidad percibida:** Se observó un nivel menor de utilidad percibida en la TL (M=5,3; SD=1,57) en comparación con la TR (M=5,8; SD=1,35). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 60).

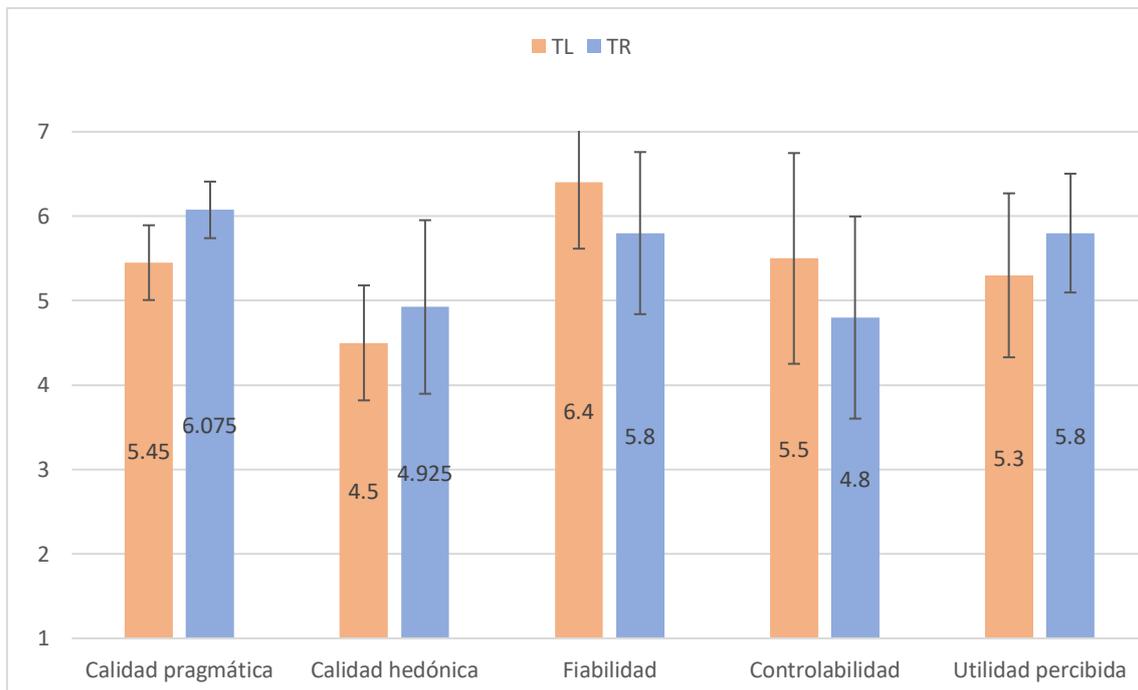


Figura 53: Resultados comparativos del cuestionario para el procedimiento LR.

Tabla 60: Resultados de la prueba T-test para determinar si existen diferencias significativas entre tareas en el procedimiento LR.

Calidad pragmática	Calidad hedónica	Fiabilidad	Controlabilidad	Utilidad percibida
0,0417	0,5093	0,3558	0,4380	0,4256
Diferencia significativa	No hay diferencia significativa			

Además, en la Figura 54 se presentan los resultados del procedimiento LR desagregados por género.

- **CP:** Tanto en hombres como en mujeres se observa menor valoración en la CP de la TL (Mujeres: M=5,3; SD=0,99; Hombres: M= 5,6; SD=0,33) que en la TR (Mujeres: M=6,1; SD=0,28; Hombres: M= 6,05; SD=0,76). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 61).

- **CH:** Tanto en hombres como en mujeres se observa menor valoración en la CH de la TL (Mujeres: M=4,4; SD=0,96; Hombres: M= 4,6; SD=1,33) que en la TR (Mujeres: M=5,15; SD=1,22; Hombres: M= 4,7; SD=2,14). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 61).
- **Fiabilidad:** En el caso de las mujeres, se observa mayor valoración en la TL (M=7; SD=0) que en la TR (M=5,6; SD=2,07). Al contrario, en los hombres se observa menor valoración de fiabilidad en la TL (M=5,8; SD=1,64) que en la TR (M=6; SD=1). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 61).
- **Controlabilidad:** Tanto en hombres como en mujeres, se observa un nivel mayor de controlabilidad en la TL (Mujeres: M=5,6; SD=2,6; Hombres: M=5,4; SD=1,51) en comparación con la TR (Mujeres: M=4,4; SD=2,3; Hombres: M=5,2; SD=1,64). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 61).
- **Utilidad percibida:** Tanto en hombres como en mujeres se observa un nivel menor de valoración en la utilidad percibida de la TL (Mujeres: M=4,6; SD=1,94; Hombres: M=6; SD=0,7) en comparación con la TR (Mujeres: M=5; SD=1; Hombres: M=6,6; SD=0,55). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 61).

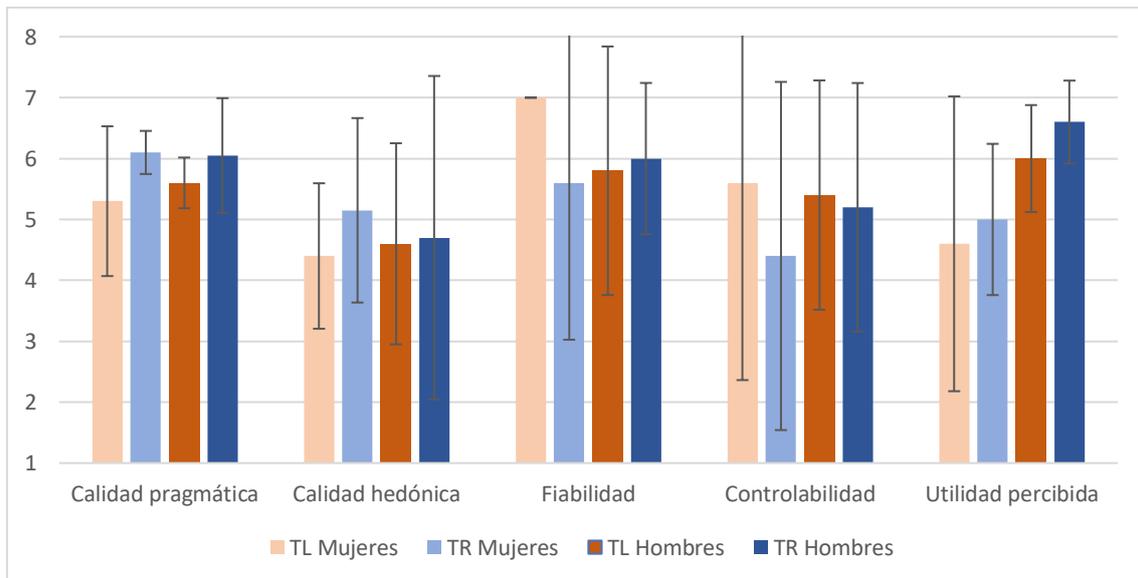


Figura 54: Resultados comparativos del cuestionario en el procedimiento LR desagregados por género.

Tabla 61: Resultados de la prueba T-test para determinar si existen diferencias significativas en las tareas desagregando por género en el Procedimiento LR.

	Calidad pragmática	Calidad hedónica	Fiabilidad	Controlabilidad	Utilidad percibida
Mujeres	0,1132 No hay diferencias significativas	0,8616 No hay diferencias significativas	0,2056 No hay diferencias significativas	0,7901 No hay diferencias significativas	0,9860 No hay diferencias significativas
Hombres	0,0860 No hay diferencias significativas	0,4902 No hay diferencias significativas	0,2825 No hay diferencias significativas	0,2062 No hay diferencias significativas	0,0753 No hay diferencia significativa

La Figura 55 muestra los resultados obtenidos en el procedimiento RL, que se describen a continuación:

- **CP:** Se observó un nivel de CP mayor en las personas participantes en la TR (M=5,97; SD=0,86) en comparación con la TL (M=4,35; SD=0,65). Esta diferencia se considera estadísticamente significativa (Tabla 62).
- **CH:** Se observó un nivel de CH mayor en las personas participantes en la TR (M=4,72; SD=1,72) en comparación con la TL (M=3,4; SD=1,28). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 62).
- **Fiabilidad:** Se observó un nivel menor de fiabilidad en la TR (M=5,9; SD=1,45) en comparación con la TL (M=5,8; SD=0,79). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 62).
- **Controlabilidad:** Se observó un nivel mayor de controlabilidad en la TL (M=6,4; SD=0,84) en comparación con la TR (M=6,1; SD=0,74). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 62).
- **Utilidad percibida:** Se observó un nivel mayor de utilidad percibida en la TR (M=5,6; SD=1,78) en comparación con la TL (M=3,3; SD=1,7). Esta diferencia no se considera estadísticamente significativa (Tabla 62).

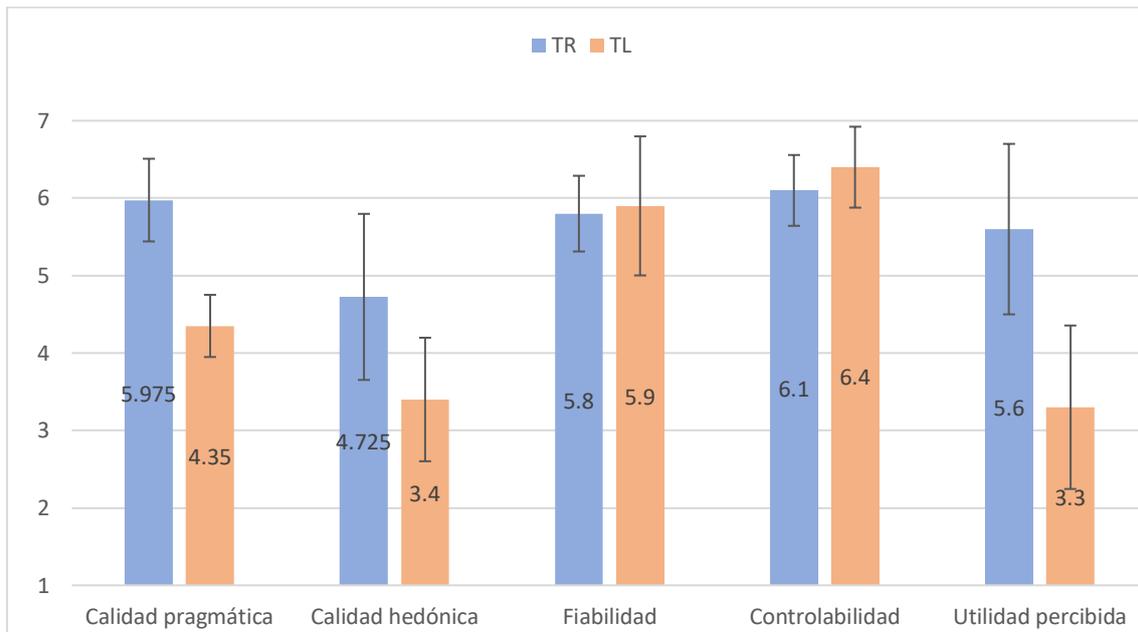


Figura 55: Resultados comparativos del cuestionario para el procedimiento RL en el caso de estudio 1.

Tabla 62: Resultados de la prueba T-test para determinar si existen diferencias significativas en el procedimiento RL.

Calidad pragmática	Calidad hedónica	Fiabilidad	Controlabilidad	Utilidad percibida
0,0002	0,0691	0,8508	0,4085	0,0085
Diferencia significativa	No hay diferencias significativas	No hay diferencias significativas	No hay diferencias significativas	Diferencia significativa

Además, en la Figura 56 se presentan los resultados del procedimiento LR desagregados por género, que se describen a continuación:

- **CP:** Tanto en hombres como en mujeres se observa mayor valoración en la TR (Mujeres: $M=6,25$; $SD=0,35$; Hombres: $M=5,7$; $SD=1,16$) que en la TL (Mujeres: $M=4,2$; $SD=0,74$; Hombres: $M=4,5$; $SD=0,59$). En el caso de las mujeres la diferencia se considera estadísticamente significativa, pero en el caso de los hombres no (Tabla 63).
- **CH:** Tanto en hombres como en mujeres se observa menor valoración en la CH de la TL (Mujeres: $M=4,85$; $SD=1,11$; Hombres: $M=4,6$; $SD=2,34$) que en la TR (Mujeres: $M=3,15$; $SD=0,68$; Hombres: $M=3,65$; $SD=1,76$). En el caso de las mujeres la diferencia se considera estadísticamente significativa, pero en el caso de los hombres no (Tabla 63).
- **Fiabilidad:** En el caso de las mujeres, se observa mayor valoración en la TR ($M=5,6$; $SD=0,55$) que en la TL ($M=5,2$; $SD=1,79$). Al contrario, en los hombres se observa menor valoración de fiabilidad en la TL ($M=6,6$; $SD=0,55$) que en la TR ($M=6$; $SD=1$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 63).
- **Controlabilidad:** En las mujeres se observa la misma valoración en la TR ($M=6$; $SD=0,7$) que en la TL ($M=6$; $SD=1$). En los hombres se observa menor percepción de controlabilidad en la TR ($M=6,2$; $SD=0,84$) que en la TL ($M=6,8$; $SD=0,45$). Estas diferencias no se consideran estadísticamente significativas (Tabla 63).
- **Utilidad percibida:** Tanto en hombres como en mujeres se observa un nivel menor de valoración en la utilidad percibida de la TL (Mujeres: $M=4,2$; $SD=1,3$; Hombres: $M=2,4$; $SD=1,6$) en comparación con la TR (Mujeres: $M=6$; $SD=0,7$; Hombres: $M=5,2$; $SD=2,5$). En el caso de las mujeres, la diferencia se considera estadísticamente significativa, aunque en el caso de los hombres no (Tabla 63).

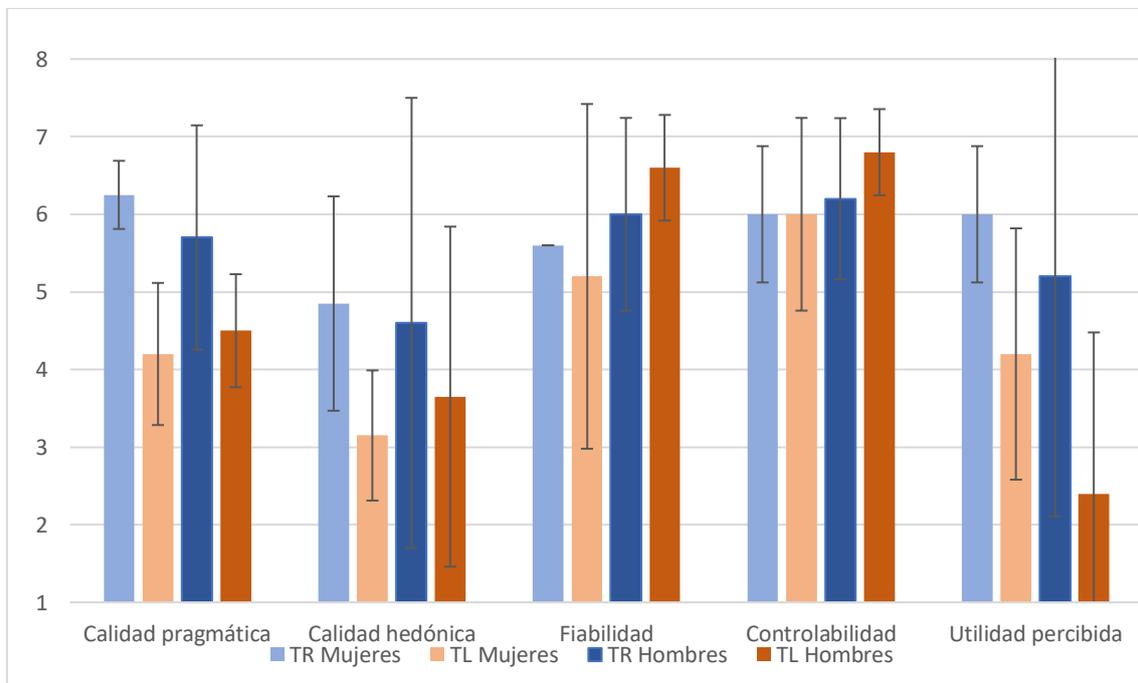


Figura 56: Resultados comparativos del cuestionario en el procedimiento RL desagregados por género en el caso de estudio 1.

Tabla 63: Resultados de la prueba T-test para determinar si existen diferencias significativas en las tareas desagregando por género en el Procedimiento RL.

	Calidad pragmática	Calidad hedónica	Fiabilidad	Controlabilidad	Utilidad percibida
	0,0016	0,0239	0,6538	1,0000	0,0340
Mujeres	Diferencia significativa	Diferencia significativa	No hay diferencias significativas	No hay diferencias significativas	Diferencia significativa
	0,0860	0,4902	0,2825	0,2062	0,0753
Hombres	No hay diferencias significativas	No hay diferencia significativa			

6.2.3.4.1 Análisis estadístico del cuestionario

Los datos recopilados a través del cuestionario pueden ser analizados mediante métodos estadísticos inferenciales, como la prueba t, con el objetivo de determinar si existen diferencias significativas entre las medias de dos grupos claramente definidos dentro de la muestra de usuarios. Específicamente, la muestra de usuarios se ha dividido en subgrupos considerando variables como el género (hombres/mujeres), el tipo de procedimiento (procedimiento LR/procedimiento RL) y la franja de edad (\leq / $>$ 30 años).

La prueba t se utiliza para evaluar la probabilidad de que las diferencias observadas entre las dos muestras de usuarios se deban al azar. Estos valores de probabilidad, conocidos como valores p, se comparan luego con un nivel de significancia α establecido en 0,05, lo que significa que se está dispuesto a aceptar un 5% de probabilidad de que el resultado sea considerado estadísticamente significativo. Valores bajos de p indican que las diferencias en las medias de los dos grupos no son resultado del azar (ver Tabla 64).

Para empezar, en la TL se destacan diferencias significativas entre los procedimientos en tres ítems: "Aburrido-Emocionante", "Convencional/Habitual-Novedoso" y "No Valioso-Valioso". En los tres ítems, las personas participantes del procedimiento LR otorgaron valoraciones más altas en comparación con los del procedimiento RL. Específicamente, en el ítem "Aburrido-Emocionante", las personas participantes del LR le dieron una calificación promedio de 3,5, mientras que los del RL asignaron un 2. En el ítem "Convencional/Habitual-Novedoso", las personas participantes del LR lo evaluaron con un promedio de 5,2, mientras que los del RL lo valoraron con un 3,6. En cuanto al ítem "No Valioso-Valioso", las personas participantes del procedimiento LR lo calificaron con un promedio de 5,3, mientras que los del RL le asignaron un 3,3. Sin embargo, en la TR no se observan diferencias significativa entre procedimientos.

En cuanto al género, en la TL, también se destacan diferencias significativas entre géneros en dos ítems. En el ítem "Complicado-Fácil", las mujeres otorgaron una calificación promedio de 5,8, mientras que los hombres dieron un promedio de 6,7. En el ítem "Confuso-Claro", las mujeres evaluaron con un promedio de 5,7, mientras que los hombres lo hicieron con un promedio de 6,9. No se observan diferencias significativas entre géneros en la TR.

En cuando a la edad, en la TL, se observan diferencias significativas entre grupos de edad en dos ítems. En el ítem "Aburrido-Emocionante", aquellos menores de 30 años lo calificaron con un promedio de 3,33, mientras que los mayores de 30 años lo valoraron con un promedio de 1,88. En el ítem "No Interesante-Interesante", los menores de 30 años otorgaron una puntuación promedio de 5,08, mientras que los mayores de 30 años dieron un promedio de 3,12.

Por otro lado, en la TR, también se observan diferencias significativas entre grupos de edad en tres ítems. En el ítem "No Interesante-Interesante", los menores de 30 años asignaron una calificación promedio de 6,08, mientras que los mayores de 30 años dieron un promedio de 3,87. En el ítem "Convencional-Original", los menores de 30 años valoraron con un promedio de 5,58, mientras que los mayores de 30 años lo hicieron con un promedio de 3,62. En el ítem "Convencional/Habitual-Novedoso", los menores de 30 años otorgaron un promedio de 5,58, en contraste con los mayores de 30 años, que dieron un promedio de 3,75.

Tabla 64: Resultados T-test por ítem (diferencias significativas por género, procedimiento y edad) considerando $\alpha = 0,05$. Se resaltan en negrita las diferencias significativas.

Categoría	Ítem	Tarea Lenta			Tarea Rápida		
		Procedimiento LR vs. RL	Género Mujeres vs. Hombres	Edad Menores de 30 vs. Mayores de 30	Procedimiento LR vs. RL	Género Mujeres vs. Hombres	Edad Menores de 30 vs. Mayores de 30
Calidad pragmática	Obstructivo-Impulsor de apoyo	0,0784	0,6435	0,4468	0,9514	0,4100	0,7766
	Complicado-Fácil	0,6310	0,0352	1,0000	0,1780	1,0000	0,5936
	Ineficiente-Eficiente	0,1927	0,1231	0,4183	0,9836	0,2628	0,1215
	Confuso-Claro	0,3645	0,0397	0,3011	0,7253	0,3419	0,7082
Calidad hedónica	Aburrido-Emocionante	0,0165	0,3370	0,0221	0,6850	0,4196	0,0878
	No interesante-Interesante	0,5017	1,0000	0,0153	0,8790	0,6214	0,0127
	Convencional-Original	0,1330	0,2082	0,2923	0,6481	0,8178	0,0258
	Convencional/Habitual -Novedoso	0,0282	0,1512	0,2192	0,3657	0,9048	0,0406
Fiabilidad	Poco fiable - Fiable.	0,2570	0,8736	0,1525	0,6271	0,4712	0,2888
Controlabilidad	Incontrolable - Controlable	0,5272	0,6820	0,4044	0,3000	0,4926	0,8767
Utilidad percibida	No valiosa - Valiosa	0,0096	0,8209	0,5836	0,4125	0,5554	0,2173

6.2.4 Discusión y conclusiones

En esta sección, se procede a la interpretación y discusión de los resultados obtenidos en el estudio que se centró en analizar el impacto de las variables cinemáticas, como la velocidad y aceleración, en los factores humanos. El objetivo principal de esta investigación ha sido indagar cómo las variaciones en la velocidad y aceleración del robot durante la realización de una tarea afectan las respuestas y emociones humanas. Los resultados se desglosan en tres áreas clave: (i) indicadores de desempeño, (ii) indicadores fisiológicos e (iii) indicadores perceptuales.

Los hallazgos indican que tanto la velocidad como la aceleración del robot influyen en la valencia, la carga de trabajo y el *engagement* de las personas participantes. En particular, se observa que la velocidad del robot impacta directamente en el tiempo requerido para finalizar una tarea, con las personas participantes completando la TR, que involucró un robot con mayor velocidad de movimientos, en un 44% menos de tiempo en comparación con la TL.

Además, el orden en que se realizan las tareas afecta a los resultados. Las personas participantes tardaron en promedio un 6% más en completar TL cuando se realizó en primer lugar, y cometieron más errores en esa tarea. La tarea realizada en segundo lugar se completa con mejores resultados que la tarea realizada en primer lugar, independientemente del orden en que se realicen las tareas. El número de errores también se reduce en la tarea realizada en segundo lugar. La TR obtiene los mejores resultados en cuanto al tiempo necesario para completar la tarea.

Las velocidades y aceleraciones más elevadas provocaron en general una mayor carga de trabajo y un menor compromiso. Los resultados también sugieren que la duración de la interacción con el robot puede influir en las respuestas emocionales, ya que interacciones más prolongadas conducen a puntuaciones más bajas de activación e impacto.

En cuanto a los valores perceptivos, la TR se valora en general más positivamente que la TL. El orden de las tareas influye en las valoraciones, siendo mejor valorado TR en el procedimiento LR que en el procedimiento RL y peor valorado la TL en el procedimiento RL que en el procedimiento LR.

A partir de estos resultados, puede concluirse que la capacidad de aprendizaje desempeña un papel crucial a la hora de configurar la experiencia de las personas participantes con el robot. Las personas participantes fueron capaces de realizar la segunda tarea mejor y con menos errores, independientemente del orden en que se realizaron las tareas. Esto sugiere que las personas participantes aprendieron de sus experiencias y mejoraron su rendimiento a medida que interactuaban con el robot.

En cuanto a las medidas fisiológicas, se obtienen hallazgos novedosos como que las personas participantes tenían una respuesta emocional más negativa en valencia cuando el robot se movía más rápido (Tabla 65), lo que indica que variables cinemáticas como la velocidad y la aceleración pueden afectar al estado emocional humano. Además, las personas participantes tenían una mayor carga de trabajo cuando el robot se movía más rápido y presentaban un menor *Engagement*, lo que indica que una velocidad más lenta y una aceleración más suave conducen a un mayor *Engagement*. Las consecuencias en cuanto a medidas fisiológicas que tuvimos cuando utilizamos variables cinemáticas más altas conducen a una peor UX. Por otro lado, la interacción prolongada con el robot da lugar a puntuaciones de impacto más bajas. Estos resultados sugieren que las variables cinemáticas y el tiempo de interacción pueden afectar a las emociones y el rendimiento humano.

Tabla 65: Resultados generales y consecuencias en base a los indicadores fisiológicos.

Variable independiente	Variable de respuesta	Resultado
Variabes cinemáticas	Valencia	Variabes cinemáticas más altas generan una menor valencia.
	Memorización	Sin influencia significativa.
	Carga mental	Variabes cinemáticas más altas crean una mayor carga mental.
	<i>Engagement</i>	Variabes cinemáticas más altas generan un menor compromiso.
	Impacto	Sin influencia significativa.
	Activación	Sin influencia significativa.
Interacción prolongada con el robot	Valencia	Sin influencia significativa.
	Memorización	Sin influencia significativa.
	Carga mental	Sin influencia significativa.
	<i>Engagement</i>	Sin influencia significativa.
	Impacto	Un tiempo de interacción más prolongado genera un menor impacto.
	Activación	Un tiempo de interacción más prolongado genera una menor activación.

En cuanto a los valores perceptuales, las personas participantes valoraron más alto la CP y la CH cuando el robot se movía más rápido (Tabla 66). Sin embargo, la velocidad del robot también tiene un efecto negativo en las valoraciones de Fiabilidad y Controlabilidad, que se valoraron más bajas cuando el robot se movía más rápido. Por otro lado, la utilidad percibida se valoró más alto cuando el robot se movía más rápido. Estos resultados sugieren que existe una ligera contradicción en los resultados, ya que las variables cinemáticas más altas conducen a una mejor CP, CH y utilidad percibida, pero sin embargo a una peor Fiabilidad y Controlabilidad.

Tabla 66: Resultados generales y consecuencias en base a los indicadores perceptuales.

Variable independiente	Variable de respuesta	Resultado
Variabes cinemáticas	Calidad pragmática (CP)	Las variables cinemáticas más altas generan una mayor CP.
	Calidad hedónica (CH)	Las variables cinemáticas más altas generan una mayor CH.
	Fiabilidad	Las variables cinemáticas más altas generan una menor Fiabilidad.
	Controlabilidad	Las variables cinemáticas más altas generan una menor Controlabilidad.
	Utilidad percibida	Las variables cinemáticas más altas generan una mayor Utilidad percibida.

Las variables cinemáticas y el tiempo de interacción pueden influir en la experiencia de las personas participantes con el robot. Las personas participantes afirmaron sentirse más seguros con el robot cuando éste se movía más rápido y, con el tiempo, se sintieron más cómodos con la interacción. Sin embargo, la mayor velocidad también tuvo un impacto negativo en la valencia, *engagement* y carga de trabajo mental de las personas participantes, y dio lugar a puntuaciones más bajas de fiabilidad y controlabilidad. Por otro lado, la mayor velocidad dio lugar a valoraciones más altas de CP, CH y utilidad percibida; por lo tanto, las personas participantes informaron de una mejor experiencia. Esto sugiere que debería existir un equilibrio entre las variables cinemáticas y el tiempo de interacción a la hora de diseñar nuevos entornos colaborativos.

6.2.4.1 *Análisis de las diferencias entre grupos*

En esta sección, se presentan los resultados de un análisis de diferencias entre condiciones en las tareas. Se han considerado factores como el procedimiento empleado (LR o RL), el género de las personas participantes (mujeres o hombres) y su edad (menores de 30 o mayores de 30) para identificar posibles diferencias significativas en los indicadores recopilados. Los análisis se realizaron utilizando pruebas t de Student, con un nivel de significancia $\alpha = 0,05$ (Tabla 67).

Los resultados revelan diferencias significativas en el tiempo requerido para completar ambas tareas según el procedimiento utilizado. Específicamente, en la TL, las personas participantes que siguieron el procedimiento LR han necesitado más tiempo que aquellos que han utilizado el procedimiento RL. Por el contrario, en la TR, las personas participantes del procedimiento RL han sido quienes han requerido de más tiempo. Estas diferencias son coherentes con la naturaleza del procedimiento y la aprendibilidad de la propia tarea, ya que, como se menciona previamente, la tarea realizada en segundo lugar tiende a ser más eficiente.

Además, se han observado diferencias significativas en la TL en cuanto a los errores cometidos, desglosados por género. En promedio, las mujeres han cometido aproximadamente la mitad de los errores que los hombres en la TL.

En cuanto a las respuestas fisiológicas, se observan diferencias significativas solo en la TL y varían según el procedimiento empleado. Concretamente, estas diferencias se evidencian en los indicadores de activación y memorización. En el procedimiento LR, las personas participantes han registrado un promedio de 18,75 en el indicador de activación, mientras que en el procedimiento RL han obtenido un valor promedio de -9,79. Estos resultados apuntan a un patrón interesante: al interactuar por segunda vez en la misma tarea, y además con el robot moviéndose más lentamente, se observa una caída significativa en los niveles de activación. Esto sugiere que la repetición de la misma tarea, junto con la disminución de la velocidad del robot, puede llevar a una reducción en la respuesta de activación de las personas participantes.

Por otro lado, en lo que respecta a la memorización, las personas participantes del procedimiento LR también han obtenido un indicador más elevado, con un valor promedio de 30,40 en comparación con las personas participantes del procedimiento RL, que han registrado un valor promedio de 23,74. No se observan más diferencias significativas en las respuestas fisiológicas en cuanto al género o a la edad.

En lo que respecta a las respuestas perceptuales, se destacan diferencias en la TL, específicamente en la CP y la Utilidad Percibida. Las personas participantes del procedimiento LR han otorgado una calificación promedio de 5,45 en la CP, en comparación con las personas participantes del procedimiento RL, que han otorgado un promedio de 4,35. Estos resultados son coherentes con la dinámica de las interacciones previas de las personas participantes. Aquellos que han experimentado primero con un robot más rápido (procedimiento RL) podrían percibir menos favorablemente al robot que ha interactuado de manera más lenta en la TL.

Por último, se detectan diferencias significativas en la valoración de la CH en función de la edad en ambas tareas. En ambos casos, el grupo de participantes menores de 30 años ha otorgado calificaciones más altas a la CH en comparación con el grupo de participantes mayores de 30 años. En la TL, las personas participantes menores de 30 años la valoraron con un promedio de 4,98, mientras que los mayores de 30 años la han calificado con un promedio de 4,78. En la TR, las personas participantes menores de 30 años han otorgado una calificación promedio de 6,17, mientras que los mayores de 30 años la han evaluado con un promedio de 5,81.

Tabla 67: Resultados T-test por indicador recogido (diferencias significativas por procedimiento, género y edad) considerando $\alpha = 0,05$. Se resaltan en negra las diferencias significativas.

		Tarea Lenta			Tarea Rápida		
		Procedimiento	Género	Edad	Procedimiento	Género	Edad
Desempeño	Tiempo	0,0122	0,1662	0,8424	0,0238	0,1208	0,4162
	Errores	1,0000	0,0394	0,1107	0,0857	0,2188	0,8343
Activación	Activación	0,0248	0,5656	0,2553	0,9981	0,7216	0,3159
	Engagement	0,4424	0,4160	0,8153	0,4321	0,3593	0,3162
	Impacto	0,6981	0,5249	0,1293	0,1758	0,2583	0,8556
Respuesta fisiológica	Memorización	0,0147	0,6430	0,9867	0,6234	0,7029	0,9224
	Valencia	0,2514	0,4894	0,2628	0,7480	0,7416	0,9949
	Carga mental	0,5590	0,0947	0,5018	0,3033	0,2255	0,3681
	CP	0,0020	0,4572	0,6137	0,7601	0,3610	0,3166
Respuesta perceptual	CH	0,0550	0,5608	0,0307	0,7949	0,6502	0,0237
	Fiabilidad	0,4220	0,8736	0,1525	1,0000	0,4712	0,2888
	Controlabilidad	0,2167	0,6820	0,4044	0,0710	0,4926	0,8767
	Utilidad percibida	0,0137	0,8209	0,5836	0,7682	0,5554	0,2173

6.2.4.2 Influencia de las variables cinemáticas en la confianza

Las variables cinemáticas influyen de tal forma que una velocidad demasiado baja del robot genera desconfianza en las personas, pensamientos y dudas sobre la precisión e "inteligencia" del robot. Sin embargo, los resultados obtenidos de los ítems Fiabilidad y Controlabilidad sugieren que variables cinemáticas más altas conducen a una peor experiencia, aunque no necesariamente significa peor confianza y seguridad percibida. En este sentido, la influencia de las variables cinemáticas sobre la confianza y la seguridad percibida fue en ocasiones contradictoria. Mientras que una velocidad demasiado baja genera desconfianza en la precisión e inteligencia del robot, los resultados de los ítems Fiabilidad y Controlabilidad sugieren que unas variables cinemáticas más altas conducen a una peor experiencia. Esta discrepancia pone de manifiesto la necesidad de seguir investigando para comprender cuál es el equilibrio óptimo. Según Beschi et al. (2020), las emociones de la persona operaria se ven más afectadas por la duración de una trayectoria que por su velocidad y aceleración.

Por otro lado, Rubagotti et al. (2021) descubrieron que cuanto menor es la velocidad del robot, más seguro se percibe. Por tanto, una velocidad y una aceleración elevadas pueden disminuir la UX porque pueden provocar una falta de control y previsibilidad. Unas variables cinemáticas demasiado elevadas impiden a las personas anticipar el movimiento del robot y reaccionar en consecuencia. Esto provoca una sensación de inseguridad, que se traduce en una falta de confianza, lo que puede repercutir negativamente en la experiencia de la persona usuaria. Además, una velocidad y una aceleración elevadas pueden plantear problemas de seguridad. Una velocidad demasiado elevada también generaría desconfianza en las personas, haciendo aflorar emociones como la ansiedad y la cognición del riesgo, lo que reduciría considerablemente la seguridad percibida y la confianza (Hanajima et al., 2017).

Para representar visualmente esta compleja relación entre las variables cinemáticas del robot y la seguridad y fiabilidad percibidas por las personas usuarias, se introduce el concepto de la "Campana de Velocidad HRI" (Figura 57). Esta representación gráfica adopta la forma de una campana de Gauss, en la que el eje x refleja las variables cinemáticas del robot, mientras que el eje y denota la seguridad y la fiabilidad percibidas por las personas usuarias. La concepción de esta campana se basa en la investigación de Rubagotti et al. (2021), quienes señalaron que tanto una velocidad como una aceleración excesivas del robot se percibían como inseguras y poco fiables por parte de las personas usuarias. Sin embargo, los resultados

del presente estudio sugieren que una velocidad y aceleración demasiado bajas también generan desconfianza en los participantes. Por lo tanto, esta representación en forma de campana de Gauss enfatiza que tanto una velocidad y aceleración excesivamente bajas como unas configuraciones cinemáticas demasiado elevadas conducen a una disminución de la seguridad y la fiabilidad percibidas.

La elección de representar esta relación mediante una campana de Gauss se justifica debido a su idoneidad para describir fenómenos naturales y sociales caracterizados por distribuciones estadísticas. La campana de Gauss es una representación continua que capta la variabilidad en las preferencias y percepciones de las personas respecto a las variables cinemáticas del robot. Esto la convierte en una elección más apropiada que una representación escalonada o en forma de pirámide, ya que estas últimas implicarían transiciones bruscas en la percepción a medida que las variables se modifican. Dado que las relaciones entre variables en el contexto de la HRI suelen ser complejas y no lineales, la campana de Gauss ofrece una representación matizada que ilustra cómo diferentes niveles de variables cinemáticas pueden influir gradualmente en la percepción de seguridad y fiabilidad de las personas usuarias.

Es relevante destacar que la variabilidad en las preferencias individuales es un aspecto fundamental. Lo que un usuario considera como "seguro y fiable" en términos de velocidad y aceleración puede diferir significativamente de la percepción de otro usuario. La "Campana de Velocidad HRI" refleja esta diversidad de perspectivas individuales respecto a las variables cinemáticas. Cada incremento Δx en la gráfica puede representar una percepción única en cada usuario. Este concepto subraya la complejidad inherente a la comprensión y adaptación de las variables cinemáticas del robot a las preferencias y requisitos individuales de las personas usuarias.

Además, esta representación gráfica sugiere que encontrar el equilibrio óptimo en las variables cinemáticas del robot constituye un desafío intrincado y altamente individualizado. Futuras investigaciones podrían orientarse hacia el desarrollo de métodos más precisos y eficientes para determinar estas configuraciones óptimas para cada usuario. Esto podría implicar la recopilación de datos sobre las preferencias individuales y la adaptación dinámica de las variables cinemáticas del robot durante la interacción. Este enfoque de personalización tiene implicaciones significativas, ya que podría sentar las bases para el desarrollo de la robótica afectiva, donde las variables cinemáticas del robot se ajusten de manera personalizada a las preferencias y necesidades de cada usuario.

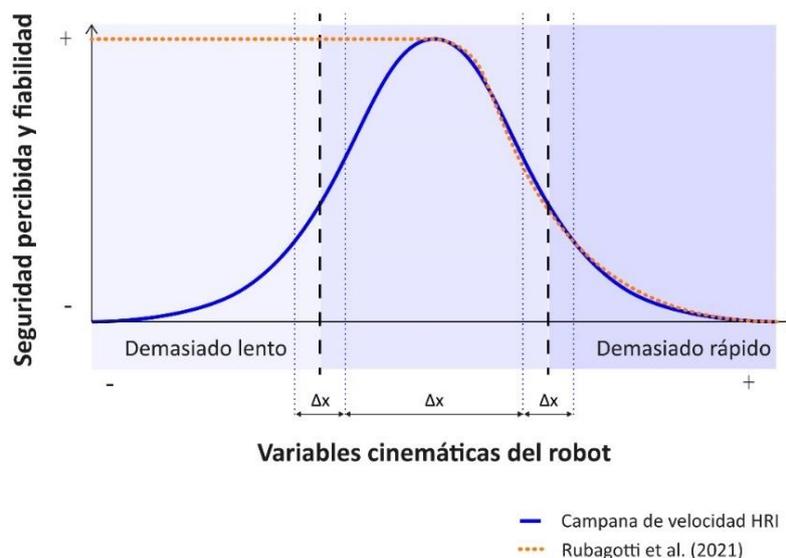


Figura 57: Representación gráfica de la "campana de velocidad en HRI".

6.2.4.3 Paradoja de la capacidad de aprendizaje

Se observa que un mayor desempeño se logra cuando el robot se mueve a una mayor velocidad y aceleración, lo que sugiere una correlación positiva entre estos factores y la eficacia percibida. Sin embargo, la secuencia en la que se presentan las tareas parece influir en las expectativas de los participantes y, en última instancia, en su capacidad para aprender y utilizar al robot de manera efectiva. En el procedimiento LR, las personas participantes mejoraron su experiencia realizando primero la TL y después la TR, lo que llevó a una mayor valoración de la utilidad percibida de la TR. Por el contrario, las personas participantes en el procedimiento RL tuvieron una experiencia peor porque primero realizaron la TR, lo que generó mayores expectativas. Cuando a continuación realizaron la TL, la experiencia empeoró considerablemente. El orden en que se realizaron las tareas en los procedimientos LR y RL tuvo un impacto significativo en el rendimiento. Los resultados mostraron que cuando las personas participantes realizaban primero la TL y después la TR, puntuaban más alto la utilidad percibida de la TR, lo que indicaba una mejor experiencia y una mayor facilidad de aprendizaje. Por otro lado, cuando las personas participantes realizaban primero la TR y después la TL, sus expectativas eran mayores, lo que provocaba una peor experiencia y una menor capacidad de aprendizaje. Esto sugiere que el orden en que se realizaron las tareas influyó enormemente en la capacidad de aprendizaje de las personas participantes y en su habilidad para utilizar el robot con eficacia.

Sin embargo, es importante reflexionar sobre el impacto de la fase de familiarización. Aunque se realizó una fase de familiarización en el preámbulo del estudio, parece que esta podría no haber sido suficientemente efectiva en la práctica. En la primera tarea, las personas participantes prácticamente se familiarizaron con el robot. Esto plantea la pregunta de si una mayor y más efectiva fase de familiarización podría haber tenido un impacto positivo en la experiencia y el aprendizaje de los participantes.

Los resultados también sugieren que el orden en que se realizaron las tareas puede haber influido en la carga mental y el *engagement*. En el procedimiento RL, las personas participantes tenían una mayor carga mental en la TR (31,10% para los hombres, 29,61% para las mujeres) y una menor carga mental en la TL (32,38% para los hombres, 28,67% para las mujeres). Del mismo modo, las personas participantes en el procedimiento LR tuvieron un mayor *engagement* en la TR (33,75% para las mujeres, 31,93% para los hombres) que en la TL (35,04% para las mujeres, 37,04% para los hombres).

Los resultados del estudio sugieren que existe un equilibrio entre capacidad de aprendizaje y experiencia. A medida que las personas participantes pasan más tiempo con el robot, mejora su rendimiento y aumenta la capacidad de aprendizaje. Sin embargo, el estudio también revela que un mayor tiempo de interacción conduce a un descenso de la activación y el impacto de las personas participantes, lo que crea una "paradoja de aprendibilidad" en la Figura 58. Esta paradoja pone de manifiesto la necesidad de seguir investigando para encontrar el equilibrio óptimo entre rendimiento y experiencia en la colaboración entre personas y robots.

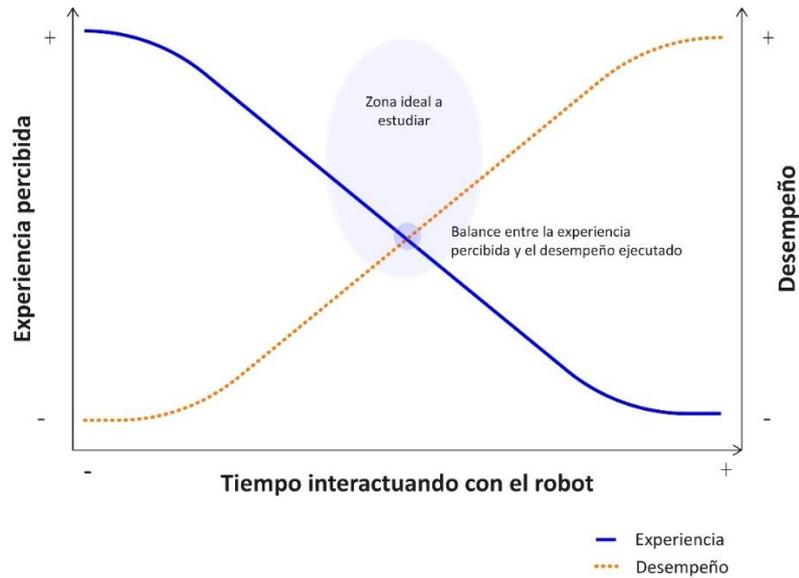


Figura 58: Representación gráfica de la paradoja de la aprendibilidad en la HRI.

6.2.4.4 La perspectiva de género

Se ha integrado la perspectiva de género en esta investigación para evitar sesgos, donde a veces las realidades, experiencias y expectativas de un grupo de personas (considerando a los hombres como referencia) se construyen como la norma, dando lugar a resultados parciales y no universales. Los resultados obtenidos, tan solo muestran diferencias estadísticamente significativas en los errores cometidos en el procedimiento LR, siendo las mujeres las que han cometido menor cantidad de errores. En lo que corresponde al género, no se han observado más diferencias estadísticamente significativas a la hora de interactuar con el robot.

6.3 Caso de estudio 2: Impacto de las interfaces unimodales vs. multimodales

En este caso de estudio, se tiene como objetivo evaluar el efecto de las interfaces unimodales en comparación con las interfaces multimodales en lo que respecta a los factores humanos y la percepción de las personas usuarias. Para llevar a cabo este análisis, se recurre al modelo ITPX-ROBOTS. A lo largo de las próximas secciones, se detallará minuciosamente el procedimiento empleado y se presentarán los resultados obtenidos en cada etapa de la investigación.

6.3.1 Motivación

La interacción eficaz entre las personas y los robots se erige como un elemento primordial para alcanzar una colaboración armoniosa y productiva. Dentro de este marco, las interfaces son el principal canal de comunicación, ya que desempeñan un papel crucial al mediar entre las personas usuarias y los sistemas robóticos. El establecimiento de una comunicación efectiva entre personas y robots a través de interfaces adecuadas se revela como un factor crítico para el éxito de la colaboración. En aras de lograrlo, se hace imperativo definir tanto las interacciones deseadas entre personas y robots como el propósito subyacente de los intercambios de información. Estos componentes se hallan en gran medida delineados por el alcance de la aplicación y las funciones específicas de las personas y los robots, y deben ser adaptados de manera contextual para asegurar su eficacia.

Uno de los aspectos relevantes que incide de manera significativa en la UX en la HRI radica en el *feedback* proporcionado a las personas usuarias a través de la interfaz (Cohen et al., 2018). Este *feedback* se presenta como un elemento esencial para que la interacción sea percibida de forma satisfactoria por las personas usuarias (Schillaci et al., 2013; Kildal et al., 2019). Cohen et al. (2018) destacan la importancia de elegir adecuadamente el modo de reacción, incluyendo factores como el canal de comunicación, la frecuencia y la intensidad del *feedback*.

Las interfaces unimodales se enfocan en un solo canal de comunicación, lo que puede simplificar la implementación y proporcionar una UX más coherente y manejable. No obstante, esta simplificación puede limitar el entendimiento y la adaptabilidad del sistema. La complementariedad de las modalidades de interfaz se convierte en un aspecto fundamental a considerar. Cada modalidad de interfaz presenta sus propias ventajas y limitaciones intrínsecas.

En este sentido, la multimodalidad de estas interfaces añade un desafío fundamental a la búsqueda de una colaboración efectiva ya que agrega un nivel de complejidad a la comunicación. Las interfaces multimodales deben ser coherentes y complementarias entre sí para que la comunicación sea efectiva. Evaluar cómo diferentes modalidades se integran y se apoyan mutuamente en la HRI es esencial para comprender su impacto en la experiencia de la persona usuaria. La efectividad de las interfaces multimodales en la HRI se mide en parte por su capacidad para permitir interacciones más naturales entre personas y robots. Evaluar si las interfaces logran este objetivo es un desafío importante.

La motivación que impulsa la realización de este caso de estudio radica en la necesidad de profundizar en la comprensión del impacto de las interfaces unimodales y multimodales en la carga mental, la memorización y la percepción global de la experiencia de las personas usuarias durante las interacciones con robots. En este sentido, se busca determinar la influencia que la elección entre interfaces unimodales y multimodales pueden tener en la experiencia de las personas usuarias.

Evaluar cómo diferentes modalidades de comunicación se integran y se apoyan mutuamente en la HRI es esencial para comprender su impacto en la experiencia de la persona usuaria. La efectividad de las interfaces multimodales en la HRI se mide en parte por su capacidad para permitir interacciones más naturales entre personas y robots.

Específicamente, en este caso de estudio, se ha optado por utilizar una interfaz unimodal basada exclusivamente en la información visual. Esta elección permite enfocarse de manera exclusiva en la evaluación de la percepción de esta modalidad en la comunicación entre la persona usuaria y el robot. Además, se ha incorporado una interfaz multimodal que combina las modalidades visual y auditiva. De este modo, se puede evaluar cómo la inclusión de la modalidad auditiva impacta en la percepción y en la interacción de la persona usuaria en contraste con la modalidad puramente visual. Esta comparativa resulta esencial para determinar si la combinación de modalidades mejora la UX y la aceptación en su desempeño en tareas específicas.

6.3.2 Metodología

El objetivo de este experimento consiste en examinar el impacto de las interfaces unimodales y multimodales en los aspectos humanos. Siguiendo el modelo ITPX-ROBOTS, se ha realizado primero una evaluación experta a través de la herramienta HEUROBOX, y a continuación un test de usuario donde se han recogido indicadores de desempeño, indicadores fisiológicos e indicadores perceptuales.

6.3.2.1 Evaluación experta a través de HEUROBOX

Para iniciar el proceso, se ha llevado a cabo una exhaustiva evaluación del sistema robótico objeto de análisis, haciendo uso de la herramienta de evaluación heurística experta a través de la versión avanzada de HEUROBOX.

6.3.2.2 Procedimiento experimental en el Test de Usuario

Este caso de estudio tiene como objetivo analizar el impacto de las interfaces en los factores humanos y en la percepción de las personas usuarias a la hora de interactuar con un robot en un entorno industrial. De acuerdo con el modelo ITPX-ROBOTS y las directrices planteadas por Apraiz, Lasa, Montagna, et al. (2023), el procedimiento experimental se ha estructurado en tres fases distintas, como se muestra en la Figura 59: (i) la fase previa a la ejecución de la tarea, (ii) la fase de ejecución de la tarea y (iii) la fase posterior a la ejecución de la tarea.

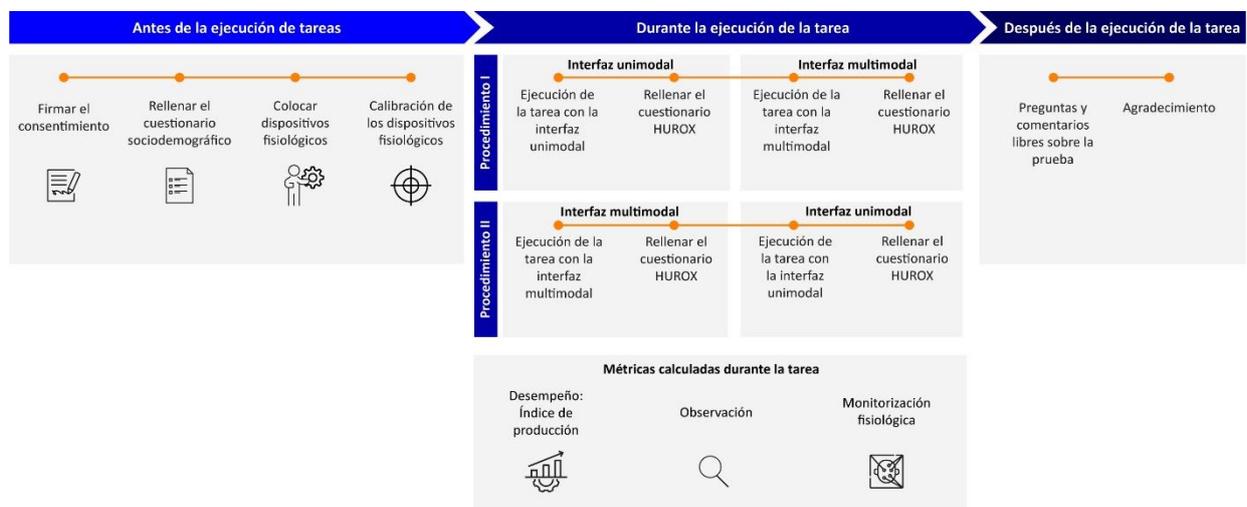


Figura 59: Procedimiento experimental en el caso de estudio 2.

6.3.2.2.1 Fase previa a la ejecución de la tarea

Inicialmente, se requirió a las personas participantes que otorgaran su consentimiento mediante la firma de un formulario informativo que detallaba los procedimientos y datos que se recopilarían durante la prueba, como grabaciones audiovisuales y la monitorización biométrica a través de electroencefalograma (EEG).

Posteriormente, se les solicitó a las personas participantes completar un cuestionario sociodemográfico. Además, se les pidió indicar su nivel de experiencia con robots siguiendo la clasificación establecida por Gervasi et al. (2022) en la Tabla 68.

Tabla 68: Niveles de experiencia previa con robot colaborativos. Fuente: Gervasi et al. (2022).

Nivel	Afirmación a seleccionar
L0	Nunca he interactuado con un robot y no los conocía hasta ahora.
L1	Nunca he interactuado con un cobot pero sé lo que son.
L2	He interactuado al menos una vez con un cobot.
L3	Ya he programado e interactuado con un cobot.

6.3.2.2.2 Durante la fase de ejecución de la tarea

Para iniciar el experimento, se colocó el EEG (Bitbrain, 2019) en las personas participantes. Conforme al *software* SennsLab (Bitbrain, 2019), se llevó a cabo un proceso de calibración de los dispositivos, siguiendo un protocolo validado proporcionado por el mismo *software* (Aliagas Ocaña, 2022). Una vez concluida la calibración del EEG, se implementaron dos procedimientos distintos:

- **En el procedimiento I**, las personas participantes (1-10) realizaron inicialmente la tarea utilizando la interfaz unimodal (exclusivamente visual). Una vez finalizada la tarea, se les solicitó completar el cuestionario HUOX. Posteriormente, repitieron la misma utilizando la interfaz multimodal (visual y voz) y, a continuación, volvieron a responder al cuestionario HUOX.
- **En el procedimiento II**, las personas participantes (11-20) llevaron a cabo inicialmente la tarea utilizando la interfaz multimodal. A continuación, completaron el cuestionario HUOX. Posteriormente, realizaron la tarea utilizando la interfaz unimodal y, una vez finalizada, volvieron a responder el cuestionario HUOX.

6.3.2.2.3 Fase posterior a la ejecución de la tarea

Posterior a la fase de interacción, se llevó a cabo una sesión de conversación abierta con las personas participantes, donde se les formuló una serie de preguntas para indagar acerca de su estado emocional y expectativas. El propósito de esta etapa era fomentar un ambiente propicio para que las personas participantes expresaran libremente sus emociones y pensamientos.

6.3.2.3 Tarea

En el marco de esta investigación, se ha concebido una tarea específica que simula el proceso de desensamblado de un frigorífico. Esta tarea está compuesta por dos objetivos interrelacionados. En primer lugar, el objetivo primordial consiste en llevar a cabo el desensamblado de un conjunto de juntas, emulando la extracción de las gomas de las puertas de las neveras. Para lograr este cometido, los participantes debían realizar diversas acciones, como la disposición de las juntas, la colaboración en la ubicación inicial del robot y la activación de su funcionamiento. El robot, por su parte, se encargaba de ejecutar la extracción de las juntas, teniendo en cuenta la ubicación de la persona, lo que significa que estaba

programado para llevar a cabo esta tarea en la dirección opuesta a la que se encontraba la persona, específicamente, hacia el lado despejado de obstáculos.

Mientras el robot se encarga de la extracción de las juntas, los participantes proceden con el desensamblado de los componentes electrónicos, compuesto por un total de 9 pasos, como se detalla en la Figura 60. A lo largo de esta tarea, se les brinda a las personas usuarias la opción de consultar la interfaz (Figura 63) con el propósito de recibir orientación paso a paso durante el proceso. Todos los componentes involucrados estaban identificados con etiquetas para garantizar que los participantes no experimentaran dificultades al identificarlos. Además, se proporciona una breve sesión de entrenamiento previa al inicio de la prueba con el fin de asegurar que los participantes pueden reconocer adecuadamente cada uno de los componentes y evitar posibles problemas en este aspecto.

Como se ilustra en la Figura 60, los dos procesos de desensamblado de componentes electrónicos se llevan a cabo de manera simultánea. En dicha figura, se proporcionan detalles paso a paso para cada uno de estos procesos. Los 9 pasos (Figura 62) que conforman el desensamblado de los componentes electrónicos son responsabilidad exclusiva de la persona, es decir, únicamente la persona usuaria se encarga de llevarlos a cabo, debiendo desmontar un total de 3 cajas. En contraste, la extracción de las juntas se realiza con la colaboración del robot. El proceso paso a paso también se detalla en la Figura 60. Como se puede observar, la persona tiene la tarea de agregar una nueva pieza y luego colocar al robot en la posición de inicio, manipulándolo físicamente, ya que algunos de sus nodos estaban diseñados específicamente para permitir este tipo de acción. Posteriormente, el participante cerraba el *gripper* del robot y lo activaba, momento en el cual el robot asume la tarea de extracción de las juntas. El robot está programado mediante técnicas de aprendizaje por refuerzo para evitar colisionar con la persona durante la ejecución de esta tarea. Como se muestra en la Figura 61, si la persona se encuentra trabajando a un lado, el robot realiza la extracción hacia el lado opuesto. A los participantes se les instruyó que debían extraer el mayor número de juntas posible mientras simultáneamente desensamblaban las cajas, pero se enfatizó que la prioridad residía en mantener al robot en funcionamiento.

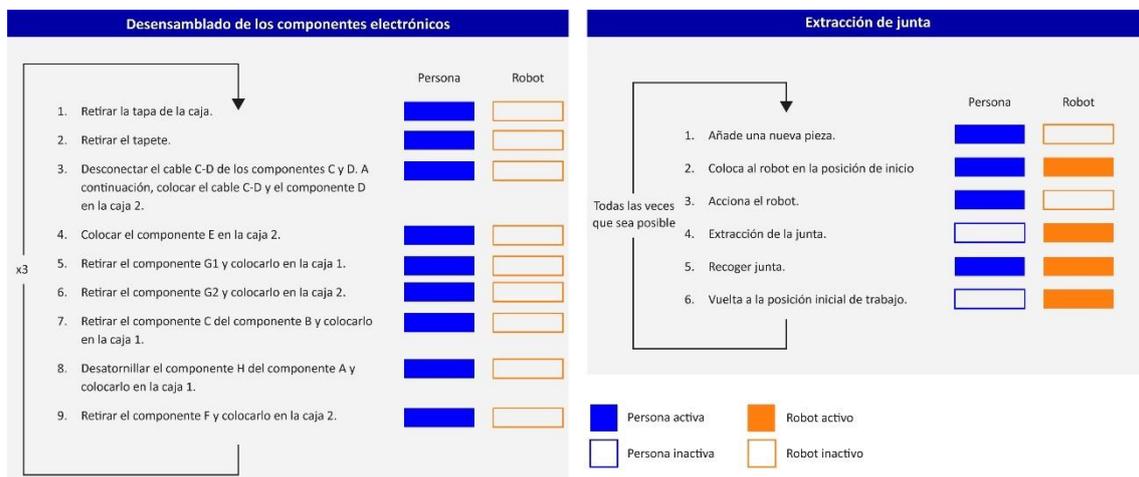


Figura 60: Secuencia de las tareas en el caso de estudio 2.

En la Figura 61, se expone la disposición empleada en el estudio. Durante el proceso en el que el participante desensamblaba la caja de los componentes electrónicos, el robot ejecutaba la extracción de la junta de la nevera en dirección opuesta a la ubicación de la persona. Es importante destacar que la interfaz se encuentra alojada en un *smartpad* móvil, lo que otorga al participante la flexibilidad de acomodarlo y posicionarlo de acuerdo a su comodidad personal.

Adicionalmente, como se puede apreciar en la Figura 61, se solicitó a los participantes que realizaran el desensamblado de dos de las cajas en el escenario A y una caja en el escenario B. Aunque no se les informó acerca del comportamiento diferenciado del robot, se les instruyó a seguir esta disposición con el propósito de que pudieran observar y evaluar cómo se sentían al respecto.

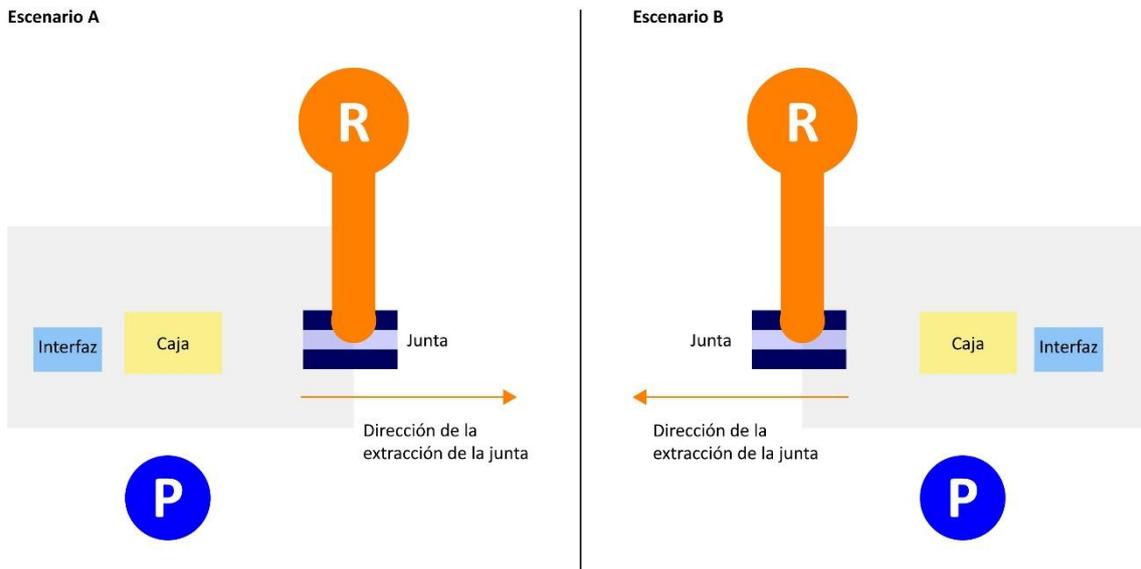


Figura 61: Set-up de la tarea en el caso de estudio 2.

En la Figura 62 se presenta el proceso paso a paso del desmontaje del componente electrónico.

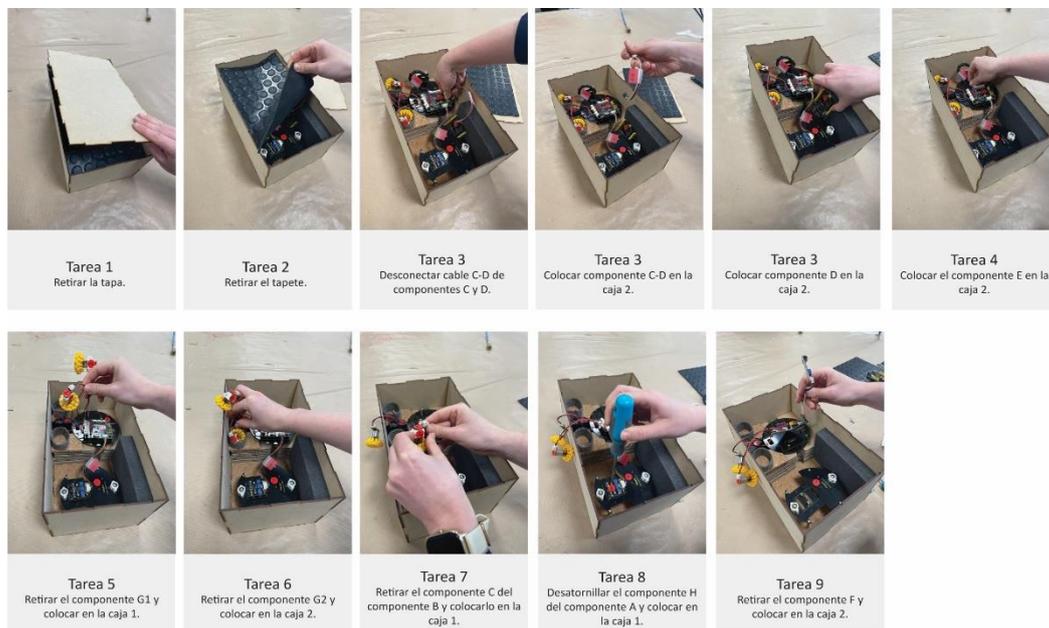


Figura 62: Paso a paso del desensamblado del componente electrónico en el caso de estudio 2.

En el proceso de desensamblado, los participantes contaron con una interfaz que les brindaba la asistencia y retroalimentación necesaria para llevar a cabo las tareas. La representación visual de esta interfaz se encuentra ilustrada en la Figura 63.

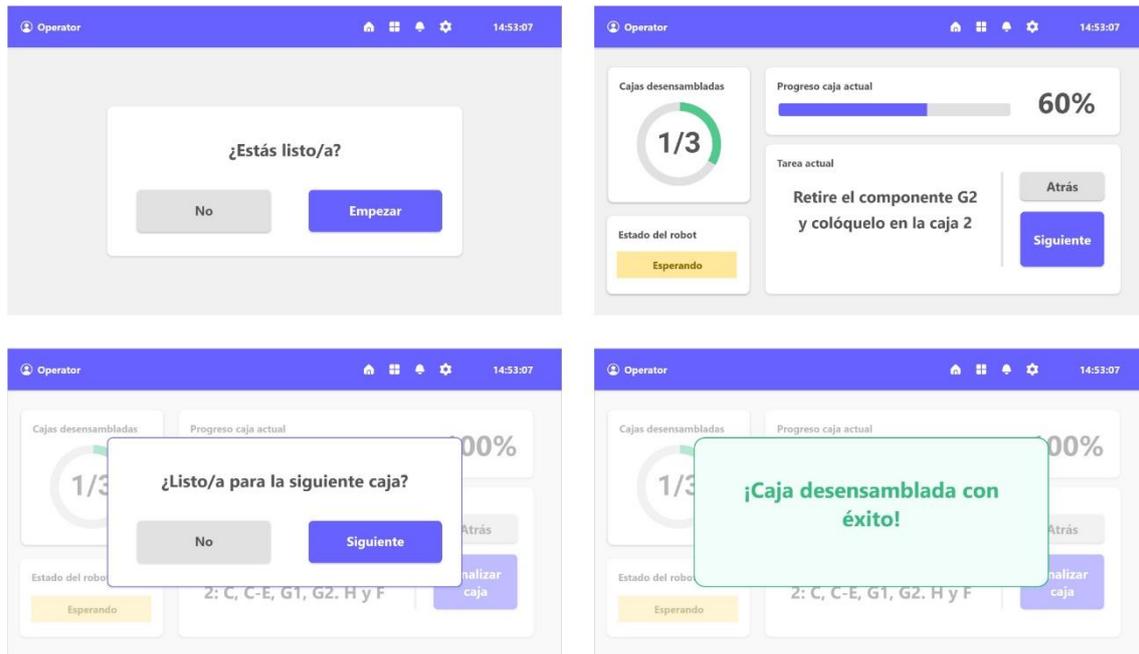


Figura 63: Ejemplos de la interfaz utilizada en el caso de estudio 2.

6.3.2.4 Indicadores del experimento

En la implementación del modelo ITPX-ROBOTS, se emplean tres categorías de indicadores: (i) indicadores de desempeño, que constituyen medidas cuantitativas y objetivas; (ii) valores fisiológicos, que son cuantitativos y objetivos; y (iii) percepciones subjetivas, que se expresan mediante valores cuantitativos y reflejan la perspectiva del usuario.

6.3.2.4.1 Indicadores de desempeño

Dado el tipo de tarea y su enfoque, se optó por registrar un único indicador de desempeño, el índice de producción. Es importante destacar que, en esta situación particular, el cálculo del tiempo necesario para llevar a cabo la tarea podría no reflejar completamente la realidad de la interacción. Esto se debe a que podría resultar en un menor número de juntas desensambladas si se completara en un tiempo más corto. Del mismo modo, el número total de juntas extraídas no proporcionaría una representación precisa, ya que podría haberse extraído un mayor número de juntas en un período de tiempo más prolongado. Para garantizar una comparación justa en términos de desempeño, se ha decidido calcular el índice de producción como el tiempo promedio necesario para extraer una junta.

6.3.2.4.2 Indicadores fisiológicos

En este caso, debido a la propia naturaleza de la tarea, las personas participantes requirieron el uso de ambas manos para ejecutarla. Como resultado, fue necesario prescindir del dispositivo Ring (Bitbrain, 2019) y, se utilizó solamente el dispositivo Diadem de Bitbrain (2019) para medir los siguientes indicadores fisiológicos: (i) valencia emocional, (ii) memorización, (iii) carga mental y (iv) *engagement*.

6.3.2.4.3 Indicadores perceptuales

Para evaluar la percepción de las personas usuarias, se empleó el cuestionario HUROY, que está compuesto por un conjunto de 7 constructos y un total de 41 ítems. Estos constructos incluyen: utilidad percibida (6 ítems), facilidad de uso percibida (9 ítems), seguridad percibida (4 ítems), controlabilidad (8 ítems), aprendibilidad (3 ítems), actitud hacia el sistema (5 ítems) y satisfacción (7 ítems). Las respuestas se registraron en una escala de Likert de 7 puntos

6.3.2.5 Procesamiento de los datos

En la Tabla 69 se resumen las variables de respuesta e independientes consideradas en el estudio. Los datos fisiológicos se recogieron con Sennslab (Bitbrain, 2019), un *software* de registro de datos multimodal. A continuación, procesaron los datos mediante Sennsmetrics (Bitbrain, 2019), una extensión de Sennslab que permite analizar los datos fisiológicos.

Tabla 69: Resumen de las variables del caso de estudio 2.

Tipo de variable	Categoría de la variable	Variable	Descripción
Variable independiente	Factores de configuración	Interfaz	Unimodal o Multimodal
Variabes de respuesta	Indicadores de desempeño	Índice de producción	Tiempo promedio necesario para llevar a cabo una extracción de una junta, considerando el número total de juntas extraídas durante un periodo de tiempo determinado.
		Utilidad percibida	El grado en que una persona cree que el uso de un sistema en particular mejoraría su desempeño laboral (Davis, 1989).
	Facilidad de uso percibida	El grado en que una persona cree que el uso de un sistema en particular sería libre de esfuerzo y fácil de aprender (Davis, 1989).	
	Seguridad percibida	El grado en que una persona percibe que un sistema en particular puede adquirir nuevos conocimientos o habilidades (Akalin et al., 2022).	
	Controlabilidad	El grado en que una persona se siente en control de una tecnología y sus acciones, y la capacidad para modificar su comportamiento según sus necesidades o preferencias (Sarker & Wells, 2003).	
	Aprendibilidad	El grado en que las personas usuarias pueden realizar con éxito una tarea cuando se encuentran con una interfaz por primera vez, así como el grado en que pueden volverse competentes en esa tarea con repeticiones crecientes de uso (Joyce, 2019).	
	Actitud e intención de uso	La evaluación general positiva o negativa de un usuario hacia una tecnología, componentes afectivos y cognitivos. Es un predictor de la intención de uso de la persona usuaria (Davis, 1989; Fishbein & Ajzen, 1977).	
	Satisfacción	El grado en que una persona está satisfecha con el uso de una tecnología, como resultado de la utilidad, facilidad de uso y confianza en su conjunto (K. C. Lee & Chung, 2009).	
	Respuestas fisiológicas	Valencia	El grado de atracción experimentado (Bitbrain, 2021).
Memorización		La intensidad de los procesos cognitivos relacionados con la formación de recuerdos futuros (Bitbrain, 2021).	
Carga mental		El uso de recursos cognitivos para llevar a	

<i>Engagement</i>	<p>cabo una tarea (Bitbrain, 2021).</p> <p>El grado de implicación o conexión entre el participante y la tarea (Bitbrain, 2021).</p>
-------------------	--

6.3.2.6 Muestra

En el experimento participaron 20 personas voluntarias (Figura 64), concretamente 10 hombres y 10 mujeres distribuidos equitativamente en ambos procedimientos (Tabla 70).

- En el procedimiento I participaron diez personas de entre 22 y 39 años ($M = 30,6$; $SD=5,9$). Respecto a la experiencia previa con robots, se encontró que 2 participantes tenían nivel L0, 3 participantes tenían nivel L1, 4 participantes tenían nivel L2, y 1 participante tenía nivel L3.
- La muestra del procedimiento II también estuvo conformada por 10 participantes, distribuidos de manera equitativa entre hombres y mujeres. Su edad varió entre 25 y 37 años ($M = 27,7$; $SD=3,6$). En relación con la experiencia previa con robots, se encontró que 4 participantes tenían nivel L2 y 6 participantes tenían nivel L3.

Tabla 70: Descripción de la muestra que ha participado en el Caso de Estudio 2.

Id	Género	Edad	Experiencia previa con robots	Procedimiento
1	Hombre	32	L0	I
2	Mujer	23	L1	I
3	Mujer	23	L2	I
4	Hombre	34	L1	I
5	Hombre	38	L2	I
6	Mujer	36	L2	I
7	Hombre	26	L2	I
8	Mujer	32	L1	I
9	Hombre	32	L3	I
10	Mujer	39	L0	I
11	Hombre	30	L2	II
12	Mujer	27	L2	II
13	Hombre	31	L2	II
14	Mujer	25	L2	II
15	Hombre	28	L3	II
16	Hombre	27	L2	II
17	Mujer	25	L2	II
18	Hombre	26	L2	II
19	Mujer	37	L2	II
20	Mujer	26	L2	II



Figura 64: Personas participantes realizando el test en el caso de estudio 2.

6.3.3 Resultados

La estructuración del análisis de datos sigue la metodología de investigación propuesta de acuerdo con el modelo ITPX-ROBOTS, dividiendo los resultados en cuatro secciones fundamentales:

1. Indicador HEUROBOX
2. Indicadores de desempeño.
3. Indicadores fisiológicos.
4. Indicadores perceptuales mediante el cuestionario HUOX.

6.3.3.1 Indicador HEUROBOX

En este contexto, y considerando que el objetivo del caso de estudio es la comparación entre un sistema con interfaz unimodal y otro con interfaz multimodal, se ha procedido a completar la sección de interfaces en dos ocasiones, resultando en la obtención de dos conjuntos de resultados.

Al analizar los resultados obtenidos, se distinguen las siguientes tendencias (consultar Tabla 71). Específicamente, la ergonomía destaca como área robusta, con un 82,6% sobre 100. Esto enfatiza una atención adecuada hacia la comodidad de la persona usuaria. Por otro lado, la seguridad y la funcionalidad exhiben un nivel del 56% y 55,56% respectivamente, sugiriendo posibles oportunidades de mejora en la protección, eficacia y capacidad operativa del sistema. Asimismo, el análisis de las interfaces en el contexto de la interfaz unimodal muestra un nivel de 62,5%, mientras que, para la interfaz multimodal, el nivel se eleva a un 65,9%. En consecuencia, el sistema unimodal presenta un puntaje total del 61,05% en la interfaz unimodal y un 61,6% en la interfaz multimodal.

Tabla 71: Resultado del indicador HEUROBOX en el caso de estudio 2.

Categoría	Interfaz Unimodal Resultado (%)	Interfaz Multimodal Resultado (%)
Seguridad	56%	56%
Ergonomía	82,6%	82,6%
Funcionalidad	55,56%	55,56%

Interfaces	62,5%	65,9%
TOTAL INDICADOR	61,05%	61,6%

6.3.3.2 Indicadores de desempeño

En lo que se refiere a los indicadores de desempeño, se han recogido, por un lado, el tiempo necesario para desensamblar cada una de las cajas, y, por ende, el tiempo total, y por otro lado, la producción en equipo junto con el robot, es decir, el número de justas extraídas.

Cabe mencionar que una extensión temporal en el proceso de desensamblado de las cajas por parte de un participante puede influir positivamente en la cantidad de juntas extraídas. Este fenómeno se atribuye al tiempo adicional otorgado debido a la necesidad de un participante de mayor tiempo para completar el desensamblado, lo que incrementa la duración de la interacción con el robot y, por consiguiente, amplifica la capacidad de extraer un mayor número de juntas. En este contexto, se ha calculado el ratio de producción, obtenido al dividir el tiempo total entre la cantidad de juntas extraídas, pues se considera una métrica más precisa y significativa para evaluar el rendimiento.

La Tabla 72 presenta los ratios de producción en el Procedimiento I, tanto en relación con la interfaz unimodal como con la interfaz multimodal. Al analizar estos datos, se pueden extraer varias conclusiones:

- **Variabilidad en los Ratios de Producción:** Se observa una variabilidad considerable en los ratios de producción entre las personas participantes, independientemente de su género. Los valores oscilan desde 0:38 hasta 1:11 en la interfaz unimodal, y desde 0:38 hasta 1:01 en la interfaz multimodal.
- **Influencia de la Interfaz:** Comparando los ratios de producción entre la interfaz unimodal y la interfaz multimodal, se nota que, en la mayoría de los casos, la interfaz multimodal muestra una mejora en los ratios de producción. Esto sugiere que la introducción de múltiples modalidades de interacción podría haber contribuido a un mejor desempeño en términos de producción.
- **Diferencias entre géneros:** Aunque no se observa una tendencia clara basada en el género, es interesante notar que tanto hombres como mujeres presentan ratios de producción variados en ambas interfaces. Esto indica que el rendimiento en la tarea no parece estar fuertemente influenciado por el género.

Tabla 72: Ratios de producción en el Procedimiento I.

Id	Género	Ratio de producción con la interfaz unimodal	Ratio de producción con la interfaz multimodal
1	Hombre	0:46	0:41
2	Mujer	0:44	0:47
3	Mujer	0:38	0:42
4	Hombre	0:41	0:40
5	Hombre	0:47	0:35
6	Mujer	0:46	0:49
7	Hombre	0:50	0:48
8	Mujer	1:11	0:37
9	Hombre	1:06	1:01
10	Mujer	1:01	0:55

Sin embargo, el resultado de la prueba T-test al comparar los ratios de producción con la interfaz unimodal frente a la interfaz multimodal es de 0,2359, lo que indica que la diferencia no es estadísticamente significativa.

La Tabla 73 presenta los ratios de producción del Procedimiento II, en el cual se invirtió el orden de uso de las interfaces multimodal y unimodal en la tarea de desensamblado colaborativo. A partir de estos resultados, se pueden derivar las siguientes conclusiones:

- **Impacto de la secuencia de la interfaz.** Al comparar los ratios de producción entre el Procedimiento II y el Procedimiento I (anteriormente analizado en la Tabla 76), se observa que la secuencia en la cual se utilizan las interfaces multimodal y unimodal afecta los resultados.
- **Variabilidad en el Rendimiento:** Al igual que en el Procedimiento I, se observa una variabilidad significativa en los ratios de producción entre las personas participantes, independientemente de su género. Los valores varían desde 0:42 hasta 1:11 en la interfaz multimodal, y desde 0:37 hasta 1:01 en la interfaz unimodal.

Tabla 73: Ratios de producción en el Procedimiento II.

Id	Género	Ratio de producción con la interfaz multimodal	Ratio de producción con la interfaz unimodal
11	Hombre	0:49	0:38
12	Mujer	0:46	0:38
13	Hombre	0:55	0:44
14	Mujer	0:45	0:44
15	Hombre	0:42	0:37
16	Hombre	0:44	0:44
17	Mujer	0:55	0:50
18	Hombre	0:56	0:42
19	Mujer	1:06	1:01
20	Mujer	1:11	0:59

Sin embargo, el resultado de la prueba T-test al comparar los ratios de producción con la interfaz unimodal frente a la interfaz multimodal es de 0,0870, lo que indica que la diferencia no es estadísticamente significativa.

Finalmente, cabe destacar el impacto de la aprendibilidad de la tarea. Los resultados sugieren que existe una mejora en los indicadores de desempeño en la tarea que se ejecuta en segundo lugar en ambos procedimientos. Esta diferencia puede atribuirse a la aprendibilidad de la tarea, es decir, la capacidad de las personas participantes para aprender y adaptarse a medida que realizan la tarea repetidamente. La ejecución previa de la tarea con una modalidad determinada puede proporcionar a las personas participantes una mayor familiaridad con el proceso, lo que les permite desarrollar estrategias más eficientes y mejorar su rendimiento en la segunda ejecución. En el caso del procedimiento I, las personas participantes experimentaron una mejora en el número de juntas extraídas y en el tiempo de ejecución al realizar la tarea con la interfaz multimodal después de haberse familiarizado con la interfaz unimodal. Lo mismo se observa en el procedimiento II, donde las personas participantes mejoraron su desempeño al pasar de la interfaz multimodal a la unimodal.

6.3.3.3 Indicadores fisiológicos

A continuación, se muestran y analizan los resultados obtenidos mediante el EEG en el Procedimiento I (Figura 65):

- **Engagement.** Se observó que las personas participantes mostraron un índice de *engagement* mayor al interactuar con la interfaz unimodal ($M = 37,07$; $SD = 3,5$) en comparación con la interfaz multimodal ($M = 33,45$; $SD = 5,32$). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 74).
- **Memorización.** Se pudo observar que las personas participantes exhibieron un índice de memorización mayor al interactuar con la interfaz unimodal ($M = 37,80$; $SD = 9,21$) en comparación con la interfaz multimodal ($M = 32,03$; $SD = 4,15$). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 74).
- **Valencia.** Los resultados mostraron que, durante la interacción con la interfaz unimodal, las personas participantes presentaron un índice de valencia más bajo ($M = -1,90$; $SD = 2,35$) en comparación con la interfaz multimodal ($M = 1,98$; $SD = 1,96$). Esta diferencia se considera estadísticamente significativa (Tabla 74), e indica que las personas participantes experimentaron emociones más negativas al utilizar la interfaz unimodal en comparación con la multimodal.
- **Carga mental.** Los resultados muestran que las personas participantes del procedimiento I, presentaron una carga mental mayor con la interfaz unimodal ($M=31,29$; $SD=2,51$) en comparación con la interfaz multimodal ($M= 28,41$; $SD=3,37$). Esta diferencia se considera estadísticamente significativa (Tabla 74).

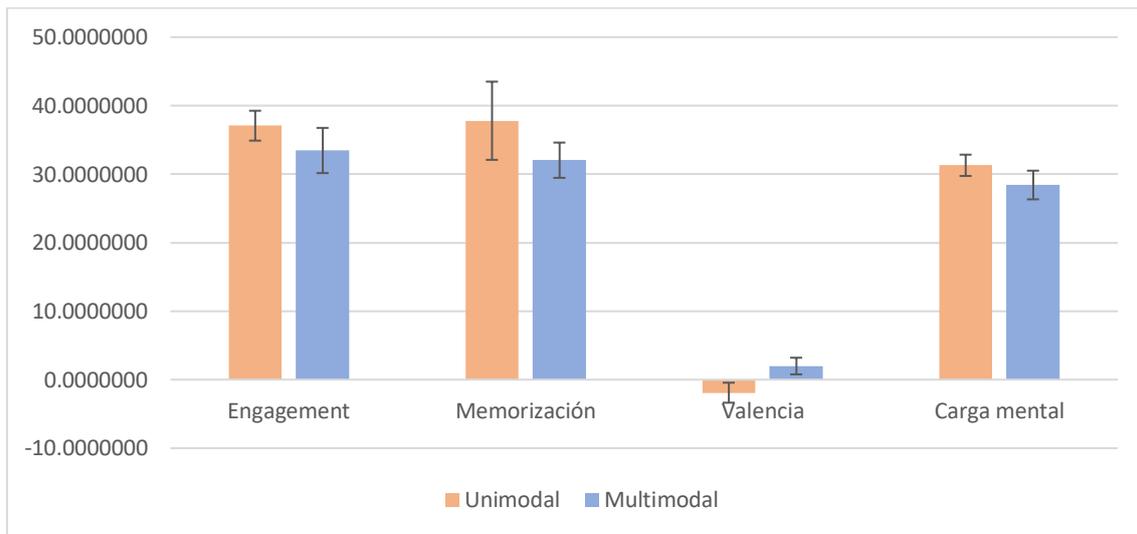


Figura 65: Resultados de los indicadores fisiológicos en el procedimiento I en el caso de estudio 2.

Tabla 74: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento I.

Engagement	Memorización	Valencia	Carga mental
0,09209	0,09547	0,00086	0,04568
No hay diferencia significativa	No hay diferencia significativa	Diferencia significativa	Diferencia significativa

Desagregando por género los datos del Procedimiento I, se obtienen los siguientes resultados (Figura 66):

- **Engagement.** Las mujeres exhibieron un índice de *engagement* más alto al interactuar con la interfaz unimodal ($M=39,57$; $SD=9,85$) en comparación con su interacción con la interfaz multimodal ($M=32,48$; $SD=7,26$). Los hombres,

por otro lado, mostraron prácticamente el mismo nivel de *engagement* tanto en la interfaz unimodal (M=34,42; SD= 0,70) como en la multimodal (M=34,43; SD=2,92). Estas diferencias no son estadísticamente significativas (Tabla 75).

- **Memorización.** Las mujeres demostraron un índice de memorización más elevado al interactuar con la interfaz unimodal (M=39,73; SD=3,13) en comparación con su interacción con la interfaz multimodal (M=33,14; SD=5,43). Por su parte, los hombres también evidenciaron una mayor memorización en la interfaz unimodal (M=36,02; SD= 9,27) que en la multimodal (M=30,92; SD=2,49). Estas diferencias no son estadísticamente significativas (Tabla 75).
- **Valencia.** Tanto las mujeres como los hombres presentan una menor valencia al interactuar con la interfaz unimodal (Mujeres: M=-0,97; SD=1,99; Hombres: M=-2,84; SD=2,52) en comparación con la interfaz multimodal (Mujeres: M=2,03; SD=1,69; Hombres: M=1,95; SD=2,40). En el caso de las mujeres esta diferencia se considera estadísticamente significativa, aunque en el caso de los hombres no (Tabla 75).
- **Carga mental.** Tanto en mujeres como en hombres, se evidencia un mayor nivel de carga mental al interactuar con la interfaz unimodal (Mujeres: M=31,93; SD=2,65; Hombres: M=30,64; SD=2,47) en comparación con la interfaz multimodal (Mujeres: M=29,76; SD=2,72; Hombres: M=27,07; SD=3,70). Estas diferencias no son estadísticamente significativas (Tabla 75).

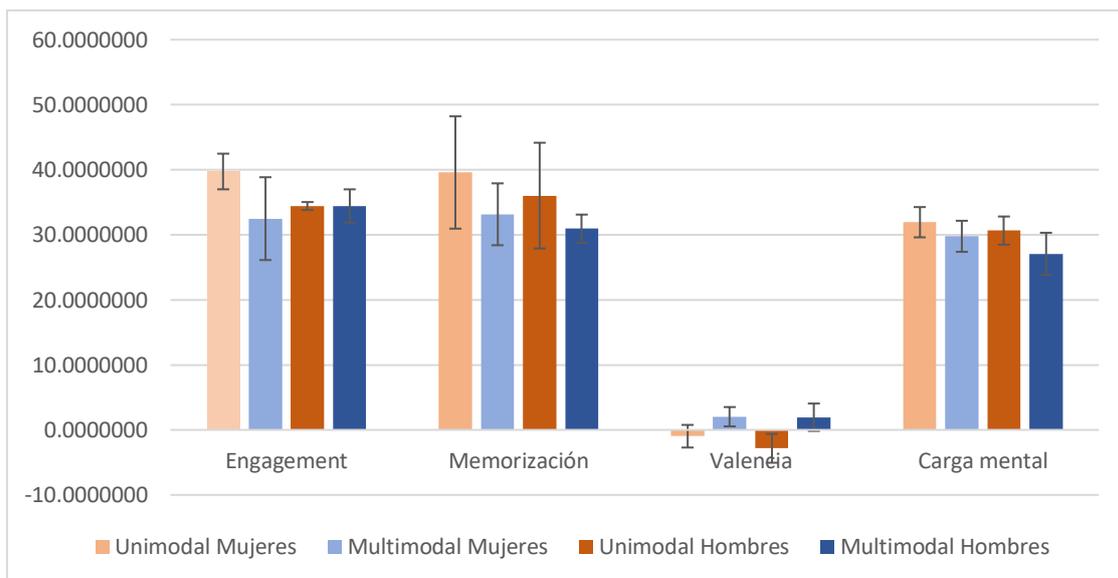


Figura 66: Resultados de los indicadores fisiológicos desagregados por género en el procedimiento I en el caso de estudio 2.

Tabla 75: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento I desagregados por género.

	<i>Engagement</i>	Memorización	Valencia	Carga mental
Mujeres	0,16584 No hay diferencia significativa	0,10467 No hay diferencia significativa	0,02039 Diferencia significativa	0,16132 No hay diferencia significativa
Hombres	0,40757 No hay diferencia significativa	0,52714 No hay diferencia significativa	0,05082 No hay diferencia significativa	0,17295 No hay diferencia significativa

A continuación, se muestran y analizan los resultados obtenidos mediante el EEG en el Procedimiento II (Figura 67):

- **Engagement.** Se observó que las personas participantes mostraron un índice de *engagement* mayor al interactuar con la interfaz multimodal ($M = 35,32$; $SD = 2,82$) en comparación con la interfaz unimodal ($M = 32,65$; $SD = 5,06$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 76).
- **Memorización.** Se pudo observar que las personas participantes exhibieron un índice de memorización menor al interactuar con la interfaz multimodal ($M = 35,43$; $SD = 7,75$) en comparación con la interfaz unimodal ($M = 38,10$; $SD = 4,58$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 76).
- **Valencia.** Los resultados mostraron que, durante la interacción con la interfaz multimodal, las personas participantes presentaron un índice de valencia más elevado ($M = 3,51$; $SD = 4,01$) en comparación con la interfaz unimodal ($M = 1,65$; $SD = 5,63$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 76)
- **Carga mental.** Las personas participantes del procedimiento II, presentaron una carga mental menor con la interfaz multimodal ($M=27,85$; $SD=3,10$) en comparación con la interfaz unimodal ($M= 30,69$; $SD=1,46$). Esta diferencia se considera estadísticamente significativa (Tabla 76).

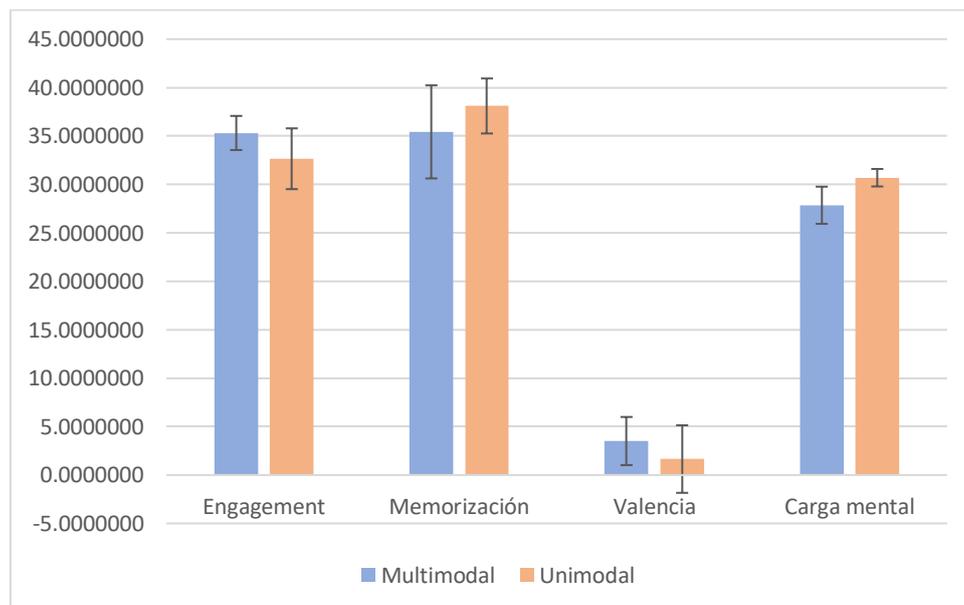


Figura 67: Resultados de los indicadores fisiológicos en el procedimiento II en el caso de estudio 2.

Tabla 76: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento II.

Engagement	Memorización	Valencia	Carga mental
0,1691	0,3627	0,4063	0,0212
No hay diferencia significativa	No hay diferencia significativa	No hay diferencia significativa	Diferencia significativa

Desagregando por género los datos del Procedimiento II, se obtiene lo siguiente (Figura 68):

- **Engagement.** Tanto mujeres como hombres presentaron un índice mayor de *engagement* al interactuar con la interfaz multimodal (Mujeres: M=35,34; SD=3,40; Hombres: M=35,29; SD=2,64) que con la interfaz unimodal (Mujeres: M=33,67; SD=4,03; Hombres: M=31,64; SD=6,23). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 77).
- **Memorización.** Tanto mujeres como hombres presentaron menor memorización al interactuar con la interfaz multimodal (Mujeres: M=37,24; SD=5,76; Hombres: M=33,61; SD=9,69) que con la interfaz unimodal (Mujeres: M=39,20; SD=4,32; Hombres: M=37,00; SD=5,06). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 77).
- **Valencia.** Tanto las mujeres como los hombres presentan una mayor valencia al interactuar con la interfaz multimodal (Mujeres: M=2,64; SD=2,10; Hombres: M=4,37; SD=5,47) en comparación con la interfaz unimodal (Mujeres: M=0,57; SD=3,65; Hombres: M=2,71; SD=7,41). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 77).
- **Carga mental.** Tanto en mujeres como en hombres, se evidencia un menor nivel de carga mental al interactuar con la interfaz multimodal (Mujeres: M=27,62; SD=2,44; Hombres: M=28,07; SD=3,94) en comparación con la interfaz unimodal (Mujeres: M=30,42; SD=1,06; Hombres: M=30,97; SD=1,86). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 77).

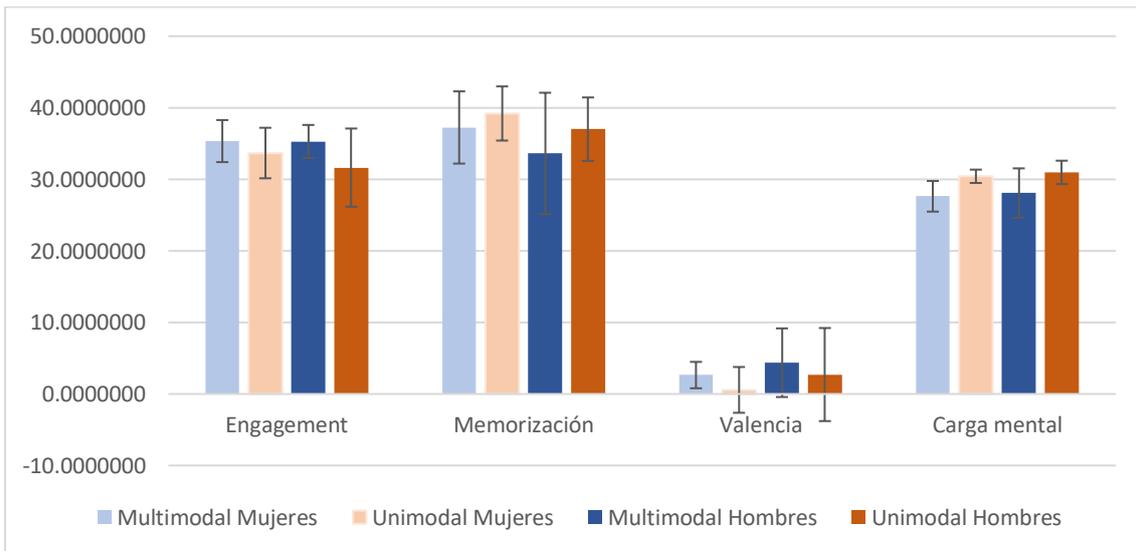


Figura 68: Resultados de los indicadores fisiológicos desagregados por género en el procedimiento II en el caso de estudio 2.

Tabla 77: Resultados de la prueba T-test para determinar si existen diferencias significativas entre el uso de las interfaces en los indicadores fisiológicos en el procedimiento II desagregados por género.

	Engagement	Memorización	Valencia	Carga mental
Mujeres	0,49640 No hay diferencia significativa	0,56164 No hay diferencia significativa	0,31175 No hay diferencia significativa	0,06125 No hay diferencia significativa
Hombres	0,27824 No hay diferencia significativa	0,51296 No hay diferencia significativa	0,69921 No hay diferencia significativa	0,19101 No hay diferencia significativa

6.3.3.4 Indicadores perceptuales

A continuación, se presentan y analizan los resultados obtenidos a través del cuestionario HUROY. Los datos correspondientes a las respuestas de las personas participantes del procedimiento I se visualizan en la Figura 69. A continuación, se exponen las conclusiones extraídas de estos resultados:

- **Utilidad percibida:** La interfaz multimodal ($M = 6,38$; $SD = 0,48$) recibió una valoración media mayor que la interfaz unimodal ($M = 5,48$; $SD = 0,95$). Esta diferencia se considera estadísticamente significativa (Tabla 78).
- **Facilidad de uso percibida:** Las personas participantes reportaron una mayor facilidad de uso en la interfaz multimodal ($M = 6,03$; $SD = 0,66$) en comparación con la interfaz unimodal ($M = 5,77$; $SD = 0,82$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 78).
- **Seguridad percibida:** Las personas participantes mostraron seguridad percibida en la interfaz multimodal ($M = 6,33$; $SD = 0,70$) en comparación con la interfaz unimodal ($M = 5,83$; $SD = 1,06$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 78).
- **Controlabilidad:** No se encontraron diferencias significativas en la percepción de controlabilidad entre ambas interfaces. Tanto la interfaz unimodal ($M = 4,19$; $SD = 1,06$) como la multimodal ($M = 4,18$; $SD = 1,18$) obtuvieron valoraciones similares en este aspecto. Consecuentemente, no se observan diferencias estadísticamente significativas (Tabla 78).
- **Aprendibilidad:** La interfaz multimodal ($M = 6,63$; $SD = 0,66$) fue percibida como más fácil de aprender en comparación con la interfaz unimodal ($M = 6,30$; $SD = 0,94$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 78).
- **Actitud e intención de uso:** No se observaron diferencias significativas en la actitud e intención de uso de las personas participantes entre ambas interfaces. Tanto la interfaz unimodal ($M = 6,02$; $SD = 1,17$) como la multimodal ($M = 6,16$; $SD = 0,81$) recibieron valoraciones similares en este aspecto. Consecuentemente, no se observan diferencias estadísticamente significativas (Tabla 78).
- **Satisfacción:** Las personas participantes reportaron una mayor satisfacción con la interfaz multimodal ($M = 6,17$; $SD = 0,72$) en comparación con la interfaz unimodal ($M = 5,71$; $SD = 1,03$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 78).

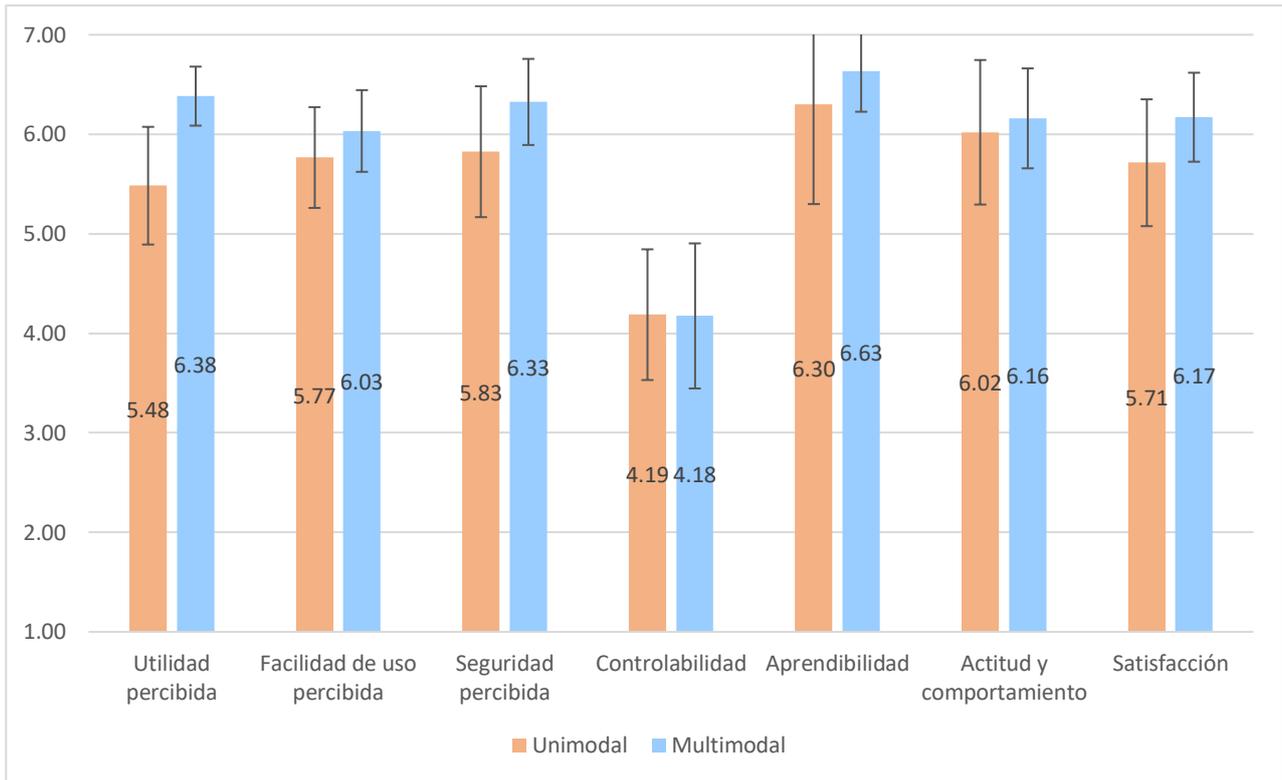


Figura 69: Resultados comparativos del cuestionario HUOX para el procedimiento I.

Además, se realizaron pruebas de T-test para determinar si las diferencias eran estadísticamente significativas en las escalas de medición evaluadas, y los resultados mostraron lo siguiente (Tabla 78):

- **Utilidad percibida:** Se encontró una diferencia significativa ($p = 0,019$) entre las evaluaciones de utilidad percibida en la interfaz unimodal y la multimodal. Esto sugiere que las personas participantes percibieron una mayor utilidad en la interfaz multimodal en comparación con la unimodal.
- **Facilidad de uso percibida, Seguridad percibida, Controlabilidad, Aprendibilidad, Actitud e intención de uso, y Satisfacción:** No se encontraron diferencias significativas en estas escalas entre la interfaz unimodal y la multimodal.

Tabla 78: Resultados de la prueba T-test para determinar si existen diferencias significativas en el procedimiento I.

Utilidad percibida	Facilidad de uso percibida	Seguridad percibida	Controlabilidad	Aprendibilidad	Actitud e intención de uso	Satisfacción
0,019	0,43	0,23	0,98	0,37	0,76	0,27
Diferencia significativa	No hay diferencias significativas					

Al realizar un análisis desagregado por género de los datos del procedimiento I (Figura 70), se pueden extraer las siguientes conclusiones:

- **Utilidad percibida.** En el caso de la interfaz unimodal, las mujeres presentaron una percepción promedio de utilidad ligeramente inferior en comparación con los hombres (Mujeres: $M = 5,40$; $SD = 0,96$; Hombres: $M = 5,57$; $SD = 1,05$). En

contraste, en la interfaz multimodal, las mujeres mostraron una mayor percepción de utilidad en comparación con los hombres, aunque esta diferencia no alcanzó significancia estadística tal y como se observa en la Tabla 79 (Mujeres: $M = 6,47$; $SD = 0,40$; Hombres: $M = 6,30$; $SD = 0,58$).

- **Facilidad de uso percibida.** En la interfaz unimodal, tanto las mujeres como los hombres otorgaron valoraciones similares en términos de facilidad de uso percibida (Mujeres: $M = 5,76$; $SD = 1,05$; Hombres: $M = 5,78$, $SD = 0,63$). Asimismo, en la interfaz multimodal, tanto las mujeres ($M = 5,82$; $SD = 0,72$) como los hombres ($M = 6,24$, $SD = 0,60$) mostraron una valoración similar de la interfaz en cuanto a su facilidad de uso percibida. Consecuentemente, no se observan diferencias estadísticamente significativas (Tabla 79).
- **Seguridad percibida.** Las mujeres otorgaron una calificación ligeramente superior a la interfaz unimodal ($M = 6,30$, $SD = 0,78$) en comparación con la interfaz multimodal ($M = 6,05$, $SD = 0,91$). Por otro lado, los hombres presentaron una diferencia más significativa, evaluando de manera menos favorable la interfaz unimodal ($M = 5,35$, $SD = 1,17$) en comparación con la interfaz multimodal ($M = 6,60$, $SD = 0,29$). Sin embargo, estas diferencias no son estadísticamente significativas (Tabla 79).
- **Controlabilidad.** Las mujeres evaluaron la controlabilidad de la interfaz unimodal ligeramente inferior en comparación con la interfaz multimodal (unimodal: $M = 4,05$, $SD = 0,73$; multimodal: $M = 4,25$, $SD = 0,47$). Por su parte, los hombres valoraron ligeramente mejor la controlabilidad de la interfaz unimodal en comparación con la interfaz multimodal (unimodal: $M = 4,33$, $SD = 1,39$; multimodal: $M = 4,1$, $SD = 1,70$). Cabe destacar que estas diferencias son muy sutiles y no resultan estadísticamente significativas (Tabla 79).
- **Aprendibilidad.** Las mujeres evaluaron de manera similar ambas interfaces (unimodal: $M = 6,80$; $SD = 0,30$; multimodal: $M = 6,80$; $SD = 0,45$). Por otro lado, los hombres valoraron la interfaz unimodal de manera inferior ($M = 5,80$; $SD = 1,12$) en comparación con la interfaz multimodal ($M = 6,47$; $SD = 0,84$). Consecuentemente, estas diferencias no son estadísticamente significativas (Tabla 79).
- **Actitud e intención de uso.** Las mujeres otorgaron una evaluación ligeramente más favorable a la interfaz unimodal ($M = 6,56$; $SD = 0,67$) en comparación con la interfaz multimodal ($M = 6,52$; $SD = 0,36$), aunque la diferencia entre ambas es mínima. Por otro lado, los hombres evaluaron de manera inferior la actitud y el comportamiento en la interfaz unimodal ($M = 5,48$; $SD = 1,38$) en comparación con la interfaz multimodal ($M = 5,80$; $SD = 1,01$). Estas diferencias no son estadísticamente significativas (Tabla 79).
- **Satisfacción.** Las mujeres reportaron una satisfacción promedio más baja con la interfaz unimodal ($M = 5,94$; $SD = 0,74$) en comparación con la interfaz multimodal ($M = 6,29$; $SD = 0,61$). Del mismo modo, los hombres también expresaron una satisfacción promedio más baja con la interfaz unimodal ($M = 5,49$; $SD = 1,30$) en comparación con la interfaz multimodal ($M = 6,06$; $SD = 0,87$). Sin embargo, estas diferencias no son estadísticamente significativas (Tabla 79).

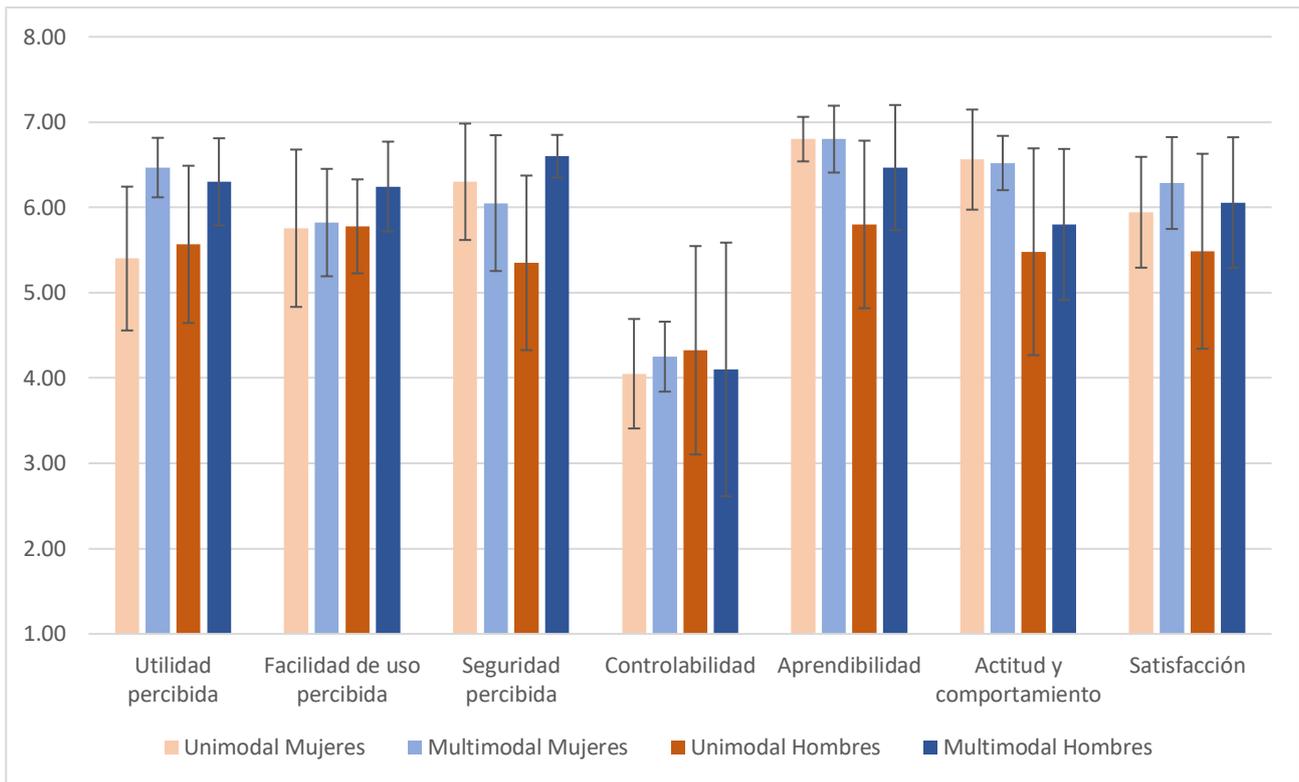


Figura 70: Resultados comparativos del cuestionario HUOX para el procedimiento I desagregados por género.

Adicionalmente, se llevaron a cabo pruebas T-test para determinar si existían diferencias significativas entre géneros en las diferentes escalas de medición. Los resultados, presentados en la Tabla 84, indican que no se encontraron diferencias estadísticamente significativas entre los géneros en ninguna de las escalas analizadas. Esto sugiere que las percepciones y evaluaciones de las interfaces unimodal y multimodal no difieren de manera significativa según el género de las personas participantes.

Tabla 79: Resultados de la prueba T-test para determinar si existen diferencias significativas entre géneros en el Procedimiento I.

	Utilidad percibida	Facilidad de uso percibida	Seguridad percibida	Controlabilidad	Aprendibilidad	Actitud e intención de uso	Satisfacción
Unimodal	0,80	0,97	0,17	0,71	0,12	0,17	0,52
	No hay diferencias significativas						
Multimodal	0,61	0,34	0,26	0,86	0,46	0,19	0,64
	No hay diferencias significativas						

Pasando al Procedimiento II, en la Figura 65 se muestran los resultados obtenidos del cuestionario HUOX (Figura 71):

- **Utilidad percibida:** Se encontró que la interfaz unimodal obtuvo una valoración media significativamente mayor ($M = 5,65$, $SD = 0,76$) en comparación con la interfaz multimodal ($M = 5,35$, $SD = 0,95$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 80)

- **Facilidad de uso percibida:** No se encontraron diferencias significativas entre ambas interfaces. Tanto la interfaz multimodal ($M = 5,52$, $SD = 0,78$) como la unimodal ($M = 5,60$, $SD = 0,80$) recibieron valoraciones similares en este aspecto. Consecuentemente, no se observan diferencias estadísticamente significativas (Tabla 80).
- **Seguridad percibida:** Se observó que la interfaz unimodal ($M = 4,45$, $SD = 0,69$) recibió una valoración media ligeramente mayor que la interfaz multimodal ($M = 4,33$, $SD = 0,53$). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 80).
- **Controlabilidad:** Se encontró que la interfaz unimodal ($M = 5,11$, $SD = 1,39$) obtuvo una valoración media mayor en comparación con la interfaz multimodal ($M = 4,60$, $SD = 1,02$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 80).
- **Aprendibilidad:** Se encontró que la interfaz unimodal ($M = 6,23$, $SD = 0,47$) fue percibida como más fácil de aprender en comparación con la interfaz multimodal ($M = 5,93$, $SD = 0,81$). Sin embargo, esta diferencia no se considera estadísticamente significativa (Tabla 80).
- **Actitud e intención de uso:** Tanto la interfaz multimodal ($M = 5,76$, $SD = 0,90$) como la unimodal ($M = 5,80$, $SD = 0,83$) recibieron valoraciones similares en este aspecto. Consecuentemente, no se observan diferencias estadísticamente significativas (Tabla 80).
- **Satisfacción:** Se encontró que la interfaz unimodal ($M = 5,16$, $SD = 0,55$) recibió una valoración media ligeramente mayor en comparación con la interfaz multimodal ($M = 4,90$, $SD = 0,67$). Sin embargo, esta diferencia no es estadísticamente significativa (Tabla 80).

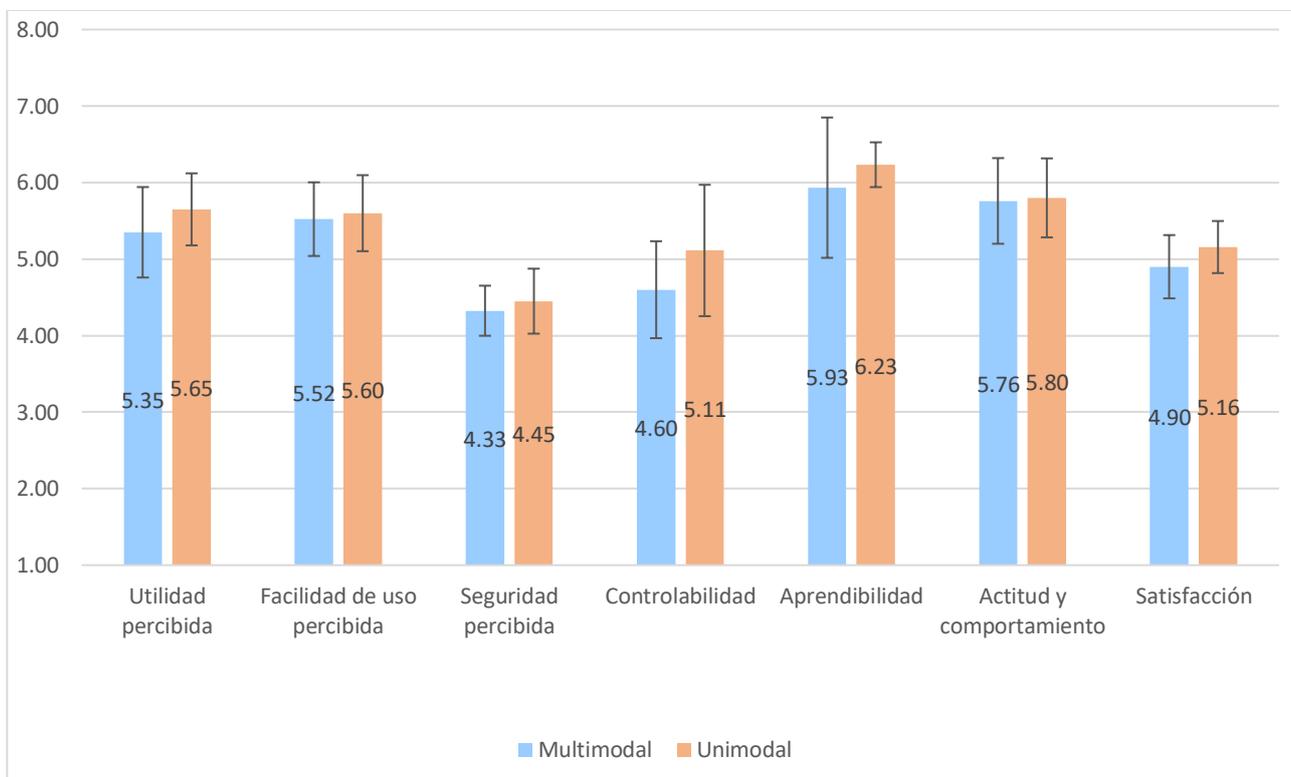


Figura 71: Resultados comparativos del cuestionario HUROY para el procedimiento II.

Se realizaron las pruebas T-test para determinar si existen diferencias significativas entre las escalas de medición evaluadas. Sin embargo, tal y como se muestra en la Tabla 85, no se identifican diferencias significativas en ninguno de los constructos.

Tabla 80: Resultados de la prueba T-test para determinar si existen diferencias significativas entre géneros en el Procedimiento II.

Utilidad percibida	Facilidad de uso percibida	Seguridad percibida	Controlabilidad	Aprendibilidad	Actitud e intención de uso	Satisfacción
0,45	0,83	0,65	0,36	0,33	0,92	0,36
No hay diferencias significativas						

Al analizar los datos del procedimiento II desagregados por género (Figura 72), se extraen las siguientes conclusiones:

- **Utilidad percibida.** En el caso de la interfaz multimodal, las mujeres presentaron una percepción promedio de utilidad ligeramente inferior en comparación con los hombres (Mujeres: $M = 5,07$; $SD = 1,13$; Hombres: $M = 5,63$; $SD = 0,75$). En la interfaz unimodal, tanto las mujeres como los hombres mostraron una mayor percepción de utilidad en comparación con la multimodal (Mujeres: $M = 5,60$; $SD = 0,35$; Hombres: $M = 5,70$; $SD = 0,32$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 81).
- **Facilidad de uso percibida.** Las mujeres mostraron valoraciones superiores tanto en la interfaz multimodal ($M = 5,76$, $SD = 0,92$) como en la interfaz unimodal ($M = 5,69$, $SD = 0,77$) en comparación con los hombres, quienes evaluaron ambas interfaces de manera similar (multimodal: $M = 5,29$, $SD = 0,61$; unimodal: $M = 5,31$, $SD = 0,73$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 81).
- **Seguridad percibida.** Las mujeres calificaron ligeramente mejor la interfaz multimodal ($M = 4,50$, $SD = 0,68$) que la interfaz unimodal ($M = 4,45$, $SD = 0,48$). Por otro lado, los hombres evaluaron ambas interfaces de manera similar (multimodal: $M = 4,15$, $SD = 0,29$; unimodal: $M = 4,15$, $SD = 0,45$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 81).
- **Controlabilidad.** Las mujeres evaluaron ligeramente peor la controlabilidad de la interfaz multimodal ($M = 4,78$; $SD = 1,08$) en comparación con la interfaz unimodal ($M = 4,89$; $SD = 1,55$), aunque la diferencia es mínima. Por otro lado, los hombres también evaluaron peor la interfaz multimodal ($M = 4,43$; $SD = 1,05$) que la interfaz unimodal ($M = 4,95$; $SD = 1,65$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 81).
- **Aprendibilidad.** Las mujeres evaluaron la aprendibilidad de la interfaz multimodal ($M = 5,53$; $SD = 0,99$) peor que la interfaz unimodal ($M = 6,63$; $SD = 0,33$). En contraste, los hombres evaluaron ambas interfaces exactamente igual ($M = 6,33$; $SD = 0,33$). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 81).
- **Actitud e intención de uso.** Las mujeres valoraron la actitud y el comportamiento hacia el sistema de manera favorable en la interfaz multimodal ($M = 5,68$; $SD = 1,12$) en comparación con la interfaz unimodal ($M = 5,48$; $SD = 0,67$). Del mismo modo, los hombres también mostraron una mejor valoración en la interfaz multimodal ($M = 5,84$; $SD = 0,75$) en comparación con la interfaz unimodal ($M = 5,68$; $SD = 0,93$).

- Satisfacción.** Las mujeres reportaron una satisfacción promedio similar tanto en la interfaz multimodal (M = 5,00; SD = 0,74) como en la interfaz unimodal (M = 5,00; SD = 0,57). Por otro lado, los hombres mostraron una satisfacción ligeramente inferior en la interfaz multimodal (M = 4,80; SD = 0,66) en comparación con la interfaz unimodal (M = 5,03; SD = 0,68). Sin embargo, estas diferencias no se consideran estadísticamente significativas (Tabla 81).

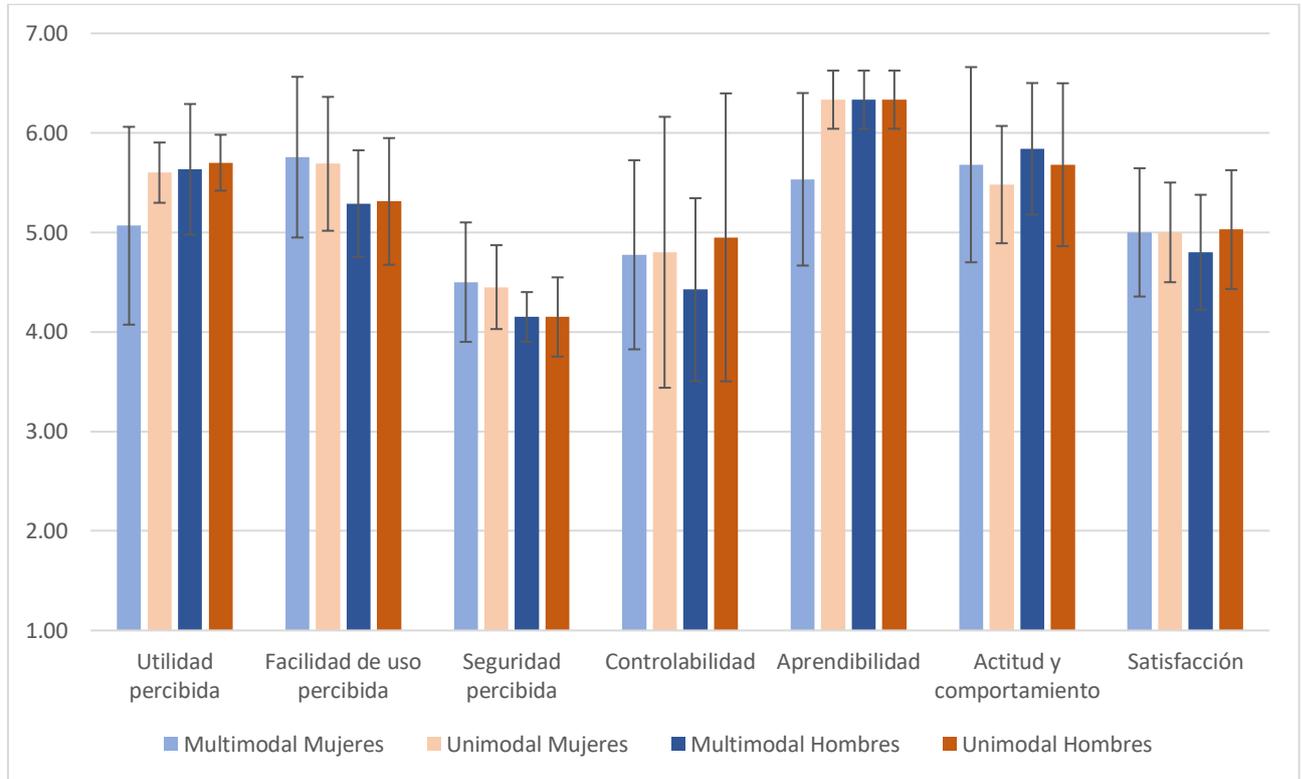


Figura 72: Resultados comparativos del cuestionario HUOX para el procedimiento II desagregados por género.

Además, se realizaron pruebas T-test para determinar si existían diferencias significativas entre géneros en las diferentes escalas de medición. Tal y como se muestra en la Tabla 81, no se identifica que las diferencias hayan sido significativas.

Tabla 81: Resultados de la prueba T-test para determinar si existen diferencias significativas entre géneros en el Procedimiento II.

	Utilidad percibida	Facilidad de uso percibida	Seguridad percibida	Controlabilidad	Aprendibilidad	Actitud e intención de uso	Satisfacción
Unimodal	0,85	0,28	0,19	0,73	0,54	0,68	0,50
	No hay diferencias significativas						
Multimodal	0,38	0,38	0,34	0,62	0,15	0,80	0,66
	No hay diferencias significativas						

A continuación, se ha analizado si existen diferencias entre procedimientos al valorar la interfaz unimodal (Figura 73). A continuación, se muestran los resultados:

- **Utilidad percibida.** En el procedimiento I, las personas participantes valoraron ligeramente inferior la utilidad percibida de la interfaz ($M = 5,48$; $SD = 0,95$) en comparación con las personas del procedimiento II ($M = 5,65$; $SD = 0,76$). Sin embargo, al realizar un análisis de T-test (Tabla 82), se considera que esta diferencia no es estadísticamente significativa.
- **Facilidad de uso percibida.** Las personas participantes del procedimiento I evaluaron la facilidad de uso percibida de la interfaz ($M = 5,77$; $SD = 0,82$) de manera ligeramente superior en comparación con las personas participantes del procedimiento II ($M = 5,60$; $SD = 0,80$). Sin embargo, al realizar la prueba T (Tabla 82), se considera que esta diferencia no alcanza significancia estadística.
- **Seguridad percibida.** Las personas participantes del Procedimiento I evaluaron la seguridad percibida de la interfaz ($M = 5,83$; $SD = 1,06$) de manera superior en comparación con las personas participantes del Procedimiento II ($M = 4,45$; $SD = 0,69$). La prueba T realizada (Tabla 82) revela que esta diferencia es estadísticamente significativa, confirmando que las personas participantes del Procedimiento I tuvieron una percepción de seguridad más favorable en comparación con los del Procedimiento II.
- **Controlabilidad.** Las personas participantes del Procedimiento I evaluaron ligeramente peor la controlabilidad de la interfaz ($M = 4,19$; $SD = 1,06$) en comparación con las personas participantes del Procedimiento II ($M = 5,11$; $SD = 1,39$). Sin embargo, la prueba T realizada (Tabla 82) no muestra una diferencia significativa entre ambos grupos, lo que indica que las discrepancias observadas podrían atribuirse al azar y no reflejan una diferencia real en la controlabilidad percibida entre los procedimientos.
- **Aprendibilidad.** Las personas participantes del Procedimiento I ($M = 6,30$; $SD = 0,94$) evaluaron ligeramente mejor la aprendibilidad de la interfaz unimodal en comparación con las personas participantes del Procedimiento II ($M = 6,26$; $SD = 0,47$). Al realizar la prueba T (Tabla 82), se determina que esta diferencia no alcanza significancia estadística.
- **Actitud e intención de uso:** Las personas que participaron en el Procedimiento I ($M = 6,02$; $SD = 1,17$) mostraron una evaluación algo más favorable en cuanto a su actitud e intención de uso de uso con la interfaz unimodal en comparación con las personas que participaron en el Procedimiento II ($M = 5,80$; $SD = 0,83$). Sin embargo, tras realizar la prueba T (Tabla 82), se concluye que esta diferencia no alcanza un nivel de significancia estadística.
- **Satisfacción:** Las personas que participaron en el Procedimiento I ($M = 5,71$; $SD = 1,03$) reportaron un nivel de satisfacción ligeramente superior con la interfaz unimodal en comparación con las personas que participaron en el Procedimiento II ($M = 5,16$; $SD = 0,55$). No obstante, tras realizar la prueba T (Tabla 82), se determina que esta diferencia no alcanza un nivel de significancia estadística.

En general, se observa una tendencia de valoraciones superiores por parte de las personas participantes del Procedimiento I, quienes han interactuado inicialmente con la interfaz unimodal sin haber utilizado ninguna otra interfaz anteriormente. Por otro lado, las personas participantes del Procedimiento II han tenido experiencia previa con la interfaz multimodal, lo que podría influir en una evaluación menos favorable de la interfaz unimodal en comparación. Esta discrepancia en las valoraciones puede atribuirse al efecto de comparación,

donde la experiencia previa con la interfaz multimodal puede afectar la percepción y evaluación de la interfaz unimodal.

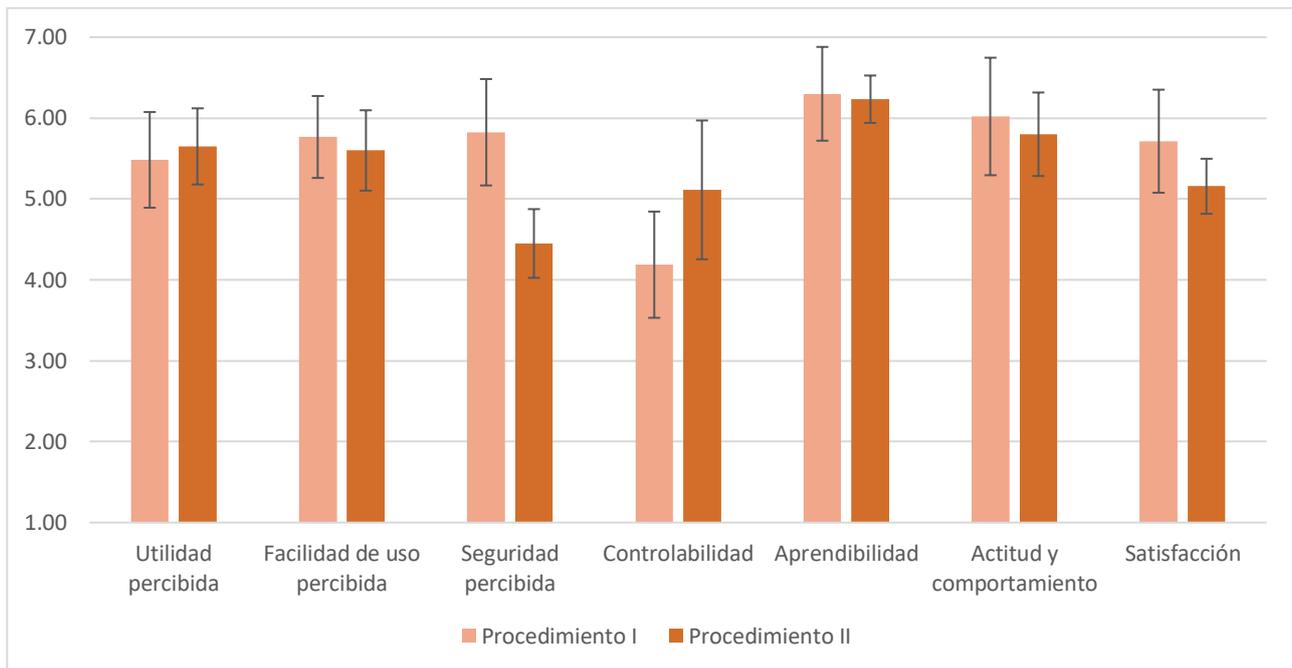


Figura 73: Resultados comparativos del cuestionario HUOX para la interfaz unimodal desagregados por procedimientos.

Además, se realizaron pruebas T-test para determinar si existían diferencias significativas entre procedimientos en las diferentes escalas de medición. Tan solo se ha identificado diferencia significativa en la seguridad percibida, siendo las personas participantes del Procedimiento I los que lo han percibido más seguro.

Tabla 82: Resultados de la prueba T-test para determinar si existen diferencias significativas entre procedimientos al valorar la interfaz Unimodal.

Utilidad percibida	Facilidad de uso percibida	Seguridad percibida	Controlabilidad	Aprendibilidad	Actitud e intención de uso	Satisfacción
0,671	0,651	0,004	0,112	0,844	0,635	0,153
No hay diferencias significativas	No hay diferencias significativas	Diferencia significativa	No hay diferencias significativas			

Tal y como se puede observar en la Figura 74, en el procedimiento I los resultados para la interfaz multimodal mostraron valores más altos en todas las escalas (excepto controlabilidad) que en el procedimiento II.

- **Utilidad percibida.** En el Procedimiento I, se observó que las personas participantes valoraron significativamente mejor la utilidad percibida de la interfaz ($M = 6,38$; $SD = 0,48$) en comparación con las personas del Procedimiento II ($M = 5,35$; $SD = 0,95$). El análisis estadístico utilizando la prueba T (Tabla 83) confirmó que esta diferencia es estadísticamente significativa ($p = 0,01$).
- **Facilidad de uso percibida.** Las personas participantes del Procedimiento I demostraron una evaluación más favorable en cuanto a la facilidad de uso percibida de la interfaz ($M = 6,03$; $SD = 0,66$) en comparación con las personas participantes del Procedimiento II ($M = 5,52$; $SD = 0,68$). No obstante, tras

realizar el análisis estadístico utilizando la prueba T (Tabla 83), se concluyó que esta diferencia no alcanza un nivel de significancia estadística.

- **Seguridad percibida.** Las personas participantes del Procedimiento I exhibieron una calificación más alta en cuanto a la seguridad percibida de la interfaz ($M = 6,33$; $SD = 0,70$) en comparación con las personas participantes del Procedimiento II ($M = 4,33$; $SD = 0,53$). Mediante la realización de la prueba T (Tabla 83), se obtuvo evidencia estadística significativa, lo cual confirma que las personas participantes del Procedimiento I experimentaron una percepción de seguridad más positiva en comparación con aquellos del Procedimiento II.
- **Controlabilidad.** Las personas participantes del Procedimiento I manifestaron una evaluación ligeramente inferior en términos de la controlabilidad de la interfaz ($M = 4,18$; $SD = 1,18$) en comparación con las personas participantes del Procedimiento II ($M = 4,60$; $SD = 1,02$). No obstante, el análisis de la prueba T realizado (Tabla 83) no reveló una diferencia significativa entre ambos grupos.
- **Aprendibilidad.** Las personas participantes del Procedimiento I demostraron una evaluación ligeramente superior en términos de la aprendibilidad de la interfaz unimodal ($M = 6,63$; $SD = 0,66$) en comparación con las personas participantes del Procedimiento II ($M = 5,93$; $SD = 0,81$). La realización de la prueba T (Tabla 83) proporcionó evidencia estadística significativa, lo que respalda la afirmación de que las personas participantes del Procedimiento I experimentaron una mayor facilidad de aprendizaje en relación con aquellos del Procedimiento II.
- **Actitud e intención de uso:** Las personas que participaron en el Procedimiento I ($M = 6,16$; $SD = 0,81$) mostraron una evaluación algo más favorable en cuanto a su actitud e intención de uso de uso con la interfaz unimodal en comparación con las personas que participaron en el Procedimiento II ($M = 5,83$; $SD = 0,81$). Sin embargo, tras realizar la prueba T (Tabla 83), se concluyó que esta diferencia no alcanza un nivel de significancia estadística.
- **Satisfacción:** Las personas que participaron en el Procedimiento I ($M = 6,17$; $SD = 0,72$) reportaron un nivel de satisfacción superior con la interfaz unimodal en comparación con las personas que participaron en el Procedimiento II ($M = 4,90$; $SD = 0,66$). La realización de la prueba T (Tabla 83) proporcionó evidencia estadística significativa, lo que respalda la afirmación de que las personas participantes del Procedimiento I experimentaron mayor satisfacción en relación con aquellos del Procedimiento II.

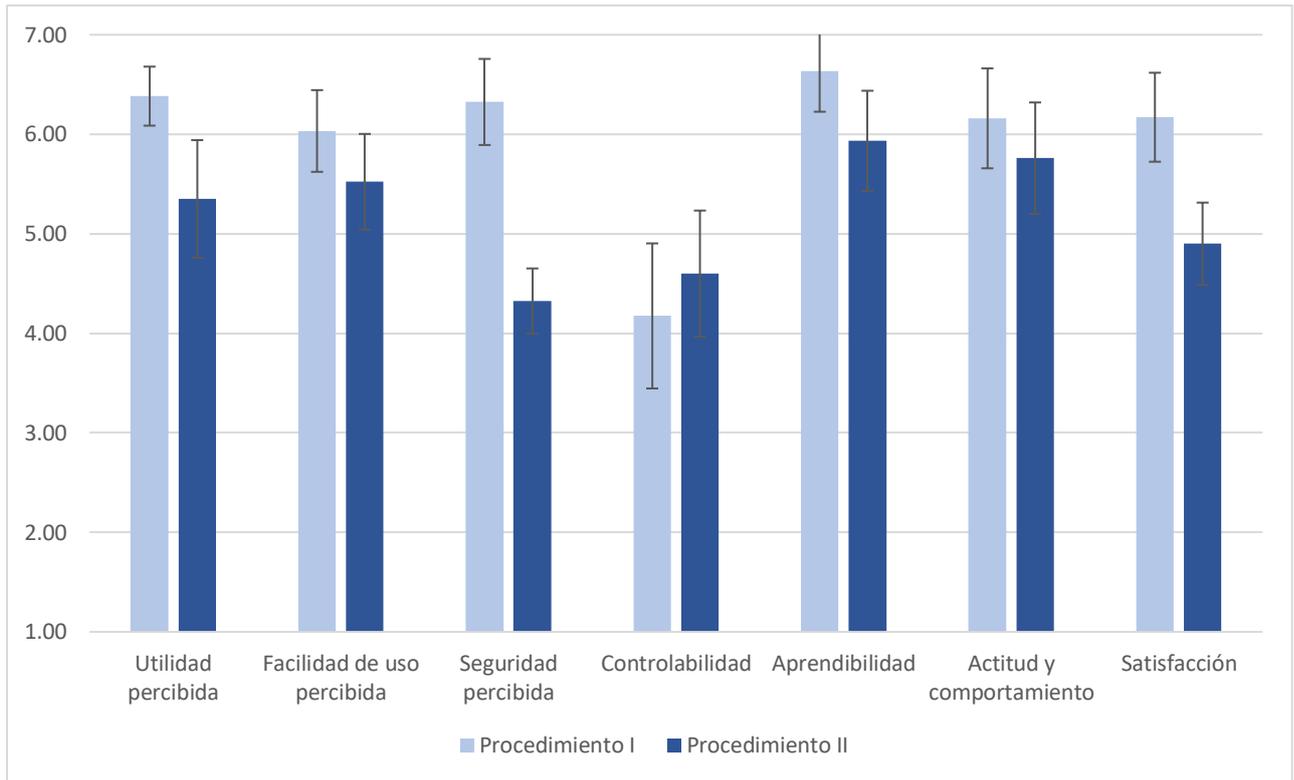


Figura 74: Resultados comparativos del cuestionario HUOX para la interfaz multimodal desagregados por procedimientos.

Al realizar pruebas T-test para determinar si las diferencias entre procedimientos son significativas (Tabla 88), se encontró que las diferencias en los constructos utilidad percibida, seguridad percibida, aprendibilidad y satisfacción son significativas.

Tabla 83: Resultados de la prueba T-test para determinar si existen diferencias significativas entre procedimientos al valorar la interfaz multimodal.

Utilidad percibida	Facilidad de uso percibida	Seguridad percibida	Controlabilidad	Aprendibilidad	Actitud e intención de uso	Satisfacción
0,01	0,13	1,53E-06	0,40	0,05	0,31	6,92E-04
Diferencia significativa	No hay diferencias significativas	Diferencia significativa	No hay diferencias significativas	Diferencia significativa	No hay diferencias significativas	Diferencia significativa

6.3.3.4.1 Análisis estadístico del cuestionario

Los datos del cuestionario pueden ser sometidos a un análisis estadístico inferencial mediante la aplicación de la prueba t, con el propósito de discernir si existen diferencias significativas entre las medias de dos grupos claramente definidos a partir de la muestra de usuarios. En específico, la muestra de usuarios ha sido segmentada considerando variables como género (hombres/mujeres), procedimiento (procedimiento I/procedimiento II), rango de edad (\leq / $>$ 30 años) y experiencia previa con robots (sin experiencia previa, aquellos que seleccionaron L0 y L1 en el cuestionario sociodemográfico / con experiencia previa, aquellos que seleccionaron L2 y L3 en el cuestionario sociodemográfico).

La prueba t evalúa la probabilidad de que los resultados de las dos muestras de usuarios consideradas se hayan manifestado de manera aleatoria. Estos valores de probabilidad (denominados valores p) son posteriormente contrastados con un umbral $\alpha =$

0,05, el cual establece una probabilidad del 5% de que el resultado sea estadísticamente significativo. Valores bajos de p indican que las diferencias en las medias de los dos grupos no se deben al azar (Tabla 84).

Mediante este análisis se puede concluir que el género no ejerce una influencia significativa en las respuestas proporcionadas por las personas usuarias. El análisis detallado de cada ítem del cuestionario revela que no se han observado diferencias estadísticamente significativas entre los géneros en ningún ítem evaluado.

En la interfaz unimodal, es importante señalar que, en dos ítems relacionados con la seguridad percibida, se observa una influencia significativa del procedimiento. Concretamente, en el ítem titulado "**Durante la tarea, me sentía ansioso/a mientras utilizaba el robot*", las personas participantes del procedimiento I proporcionaron una valoración promedio de 5,1 sobre 7 a la interfaz unimodal. Por otro lado, las personas participantes del procedimiento II otorgaron una valoración significativamente menor, con un promedio de 2,9 sobre 7 en este ítem, experimentando por tanto mayores niveles de ansiedad durante la interacción con el robot, ya que este ítem se encuentra en una escala negativa y se ha traspuesto el resultado promedio. De manera similar, en el ítem denominado "**Durante la tarea, experimenté incomodidad al utilizar el robot*", se ha registrado un patrón similar. Las personas participantes del procedimiento I indicaron un nivel de incomodidad menor, con una valoración promedio de 5,7 sobre 7. En contraste, las personas participantes del procedimiento II manifestaron experimentar una incomodidad mayor al interactuar con el robot, con una valoración promedio de 2,7 sobre 7 en este ítem.

Dentro del constructo de Controlabilidad, se destaca una disparidad significativa en el ítem titulado "**Puedo configurar el robot según mis necesidades*". Los resultados revelan una clara diferencia en las valoraciones proporcionadas por las personas participantes según el procedimiento al que estaban asignados. En el caso del procedimiento I, las personas participantes otorgaron una valoración promedio de 2,6 sobre 7 en este ítem. Las personas participantes del procedimiento II otorgaron un valor mayor de 4,8 sobre 7 de promedio.

En el constructo Actitud e intención de uso se muestra una disparidad significativa en el ítem "*Podría utilizar el robot en mi trabajo diariamente*". Las personas participantes del procedimiento I lo valoraron con un 6,6 sobre 7 de media y los del procedimiento II con un 5,7 sobre 7.

Finalmente, en el constructo de Satisfacción, se destaca una diferencia significativa en el ítem "**El robot me hace sentir incómodo/da*". Las personas participantes del procedimiento I lo valoraron con un promedio de 4,9 sobre 7, en contraste con el promedio de 2 sobre 7 asignado por las personas participantes del procedimiento II.

En la interfaz multimodal, dentro del constructo de Utilidad Percibida, se identifican tres ítems que revelan diferencias significativas en función del procedimiento aplicado. En el ítem "*El robot mejora mi rendimiento en la tarea*", se constata una disimilitud en las valoraciones proporcionadas por las personas participantes. Aquellos sometidos al procedimiento I evaluaron este aspecto con un promedio de 6,6 sobre 7, mientras que las personas participantes del procedimiento II otorgaron un promedio de 5,5 sobre 7. En el ítem "*El robot me ayuda a ser más productivo*", nuevamente se aprecia una diferencia significativa en las valoraciones entre los dos grupos de participantes. Los involucrados en el procedimiento I expresaron una valoración promedio de 6,8 sobre 7, contrastando con el promedio de 5,6 sobre 7 otorgado por las personas participantes del procedimiento II. Por último, en el ítem "*El robot me da más control sobre las tareas*", se nota un discernimiento en las valoraciones proporcionadas por las personas participantes de ambos procedimientos. Aquellos asignados al procedimiento I evaluaron este aspecto con un promedio de 5,8 sobre 7, mientras que las personas participantes del procedimiento II lo calificaron con un promedio de 4,5 sobre 7.

En el constructo de Facilidad de Uso Percibida, se destaca un ítem que exhibe una diferencia significativa entre los distintos procedimientos aplicados. En concreto, en el ítem "*El robot es amigable*", se observa una disparidad en las valoraciones otorgadas por las personas participantes. Los individuos sometidos al procedimiento I evaluaron este aspecto con un promedio de 6,8 sobre 7, mientras que aquellas personas participantes del procedimiento II asignaron un promedio de 5,3 sobre 7.

Dentro del marco del constructo de Seguridad Percibida, se identifican tres ítems que presentan diferencias significativas según el procedimiento aplicado. En particular, en el ítem "*Me siento seguro/a cuando utilizo el robot*", las personas participantes sometidos al procedimiento I asignaron un valor promedio de 1,4 sobre 7, mientras que las personas participantes del procedimiento II otorgaron un valor de 3,4 sobre 7. En relación al ítem "**Durante la tarea, me sentía ansioso/a mientras utilizaba el robot**", las personas participantes del procedimiento I otorgaron una valoración promedio de 1,7 sobre 7, en contraste con las personas participantes del procedimiento II, quienes asignaron un valor de 3,4 sobre 7. Asimismo, en el ítem "**Durante la tarea, experimenté incomodidad al utilizar el robot**", se observa una diferencia relevante en las valoraciones. Las personas participantes del procedimiento I otorgaron un valor promedio de 5,8 sobre 7, mientras que las personas participantes del procedimiento II indicaron un valor de 2,2 sobre 7.

En el constructo de Aprendibilidad, el ítem "*He adquirido destreza con el robot fácilmente*" obtuvo una valoración de 6,8 sobre 7 por parte de las personas participantes sometidos al procedimiento I, mientras que aquellos involucrados en el procedimiento II otorgaron una calificación de 5,6 sobre 7.

Dentro del marco del constructo de Satisfacción, se destacan cuatro ítems que revelan diferencias significativas. En el ítem "*Estoy satisfecho/a con el robot*", las personas participantes del procedimiento I otorgaron una calificación promedio de 6,5 sobre 7, mientras que los del procedimiento II registraron un valor de 5,6 sobre 7. En relación al ítem "*El uso del robot es armonioso*", las personas participantes del procedimiento I lo evaluaron con una puntuación de 6 sobre 7, en tanto que aquellos sometidos al procedimiento II expresaron un valor de 4,6 sobre 7. Asimismo, en el ítem "*Me gusta usar el robot*", la puntuación promedio de las personas participantes del procedimiento I asciende a 6,6 sobre 7, mientras que en el procedimiento II se registró una calificación de 5,3 sobre 7. En lo que respecta al ítem "*Me siento cómodo/da con el robot*", las personas participantes del procedimiento I asignaron una valoración de 6,6 sobre 7, en contraste con el procedimiento II, que arroja un valor de 5,5 sobre 7.

En la interfaz unimodal, dentro de la categoría de **Utilidad percibida**, se observa una diferencia significativa entre las personas participantes sin y con experiencia en la valoración del ítem "El robot me da más control sobre las tareas." Las personas participantes sin experiencia previa otorgaron una valoración promedio de 3,6 sobre 7, mientras que aquellos con experiencia previa asignaron una valoración significativamente mayor, con un promedio de 5,06 sobre 7.

En cuanto a la categoría de **Controlabilidad**, se destaca una disparidad en la valoración de dos ítems. En el ítem "Puedo configurar el robot según mis necesidades.", las personas participantes sin experiencia previa otorgaron una valoración promedio de 1,40 sobre 7, en contraste con las personas participantes con experiencia previa, que asignaron una valoración más alta, con un promedio de 4,46 sobre 7. Similarmente, en el ítem "Puedo ajustar el robot para que me ayude a realizar mis tareas.", las personas participantes sin experiencia previa valoraron este aspecto con un promedio de 1,60 sobre 7, mientras que las personas participantes con experiencia previa otorgaron una valoración más elevada, con un promedio de 4,8 sobre 7.

En la interfaz multimodal, en la categoría de Utilidad percibida, se identifican diferencias en la valoración de un ítem. En el ítem "*El robot me ayuda a ser más productivo*", las personas participantes sin experiencia previa asignaron una valoración promedio de 6,80 sobre 7, mientras que aquellos con experiencia previa le otorgaron una valoración de 6 sobre 7.

Dentro de la categoría de Seguridad percibida, se observa una disparidad en la valoración del ítem "*Me siento seguro/ra cuando utilizo el robot*". Las personas participantes sin experiencia previa asignaron una valoración promedio de 6,80 sobre 7, mientras que las personas participantes con experiencia previa le otorgaron una valoración ligeramente menor, con un promedio de 5,86 sobre 7.

En relación a la Controlabilidad, se destaca una diferencia significativa en la valoración de dos ítems. En el ítem "*Puedo configurar el robot según mis necesidades*", las personas participantes sin experiencia previa otorgaron una valoración promedio de 1,40 sobre 7, mientras que aquellos con experiencia previa asignaron una valoración más alta, con un promedio de 3,87 sobre 7. Similarmente, en el ítem "*Puedo ajustar el robot para que me ayude a realizar mis tareas*", las personas participantes sin experiencia previa valoraron este aspecto con un promedio de 1,40 sobre 7, mientras que las personas participantes con experiencia previa le asignaron una valoración más alta, con un promedio de 4,13 sobre 7.

En la categoría de Satisfacción, se identifican diferencias en la valoración del ítem "**El robot me hace sentir incómodo/da*". Las personas participantes sin experiencia previa otorgaron una valoración promedio de 6,40 sobre 7, mientras que aquellos con experiencia previa le asignaron una valoración significativamente menor, con un promedio de 3,33 sobre 7.

En la interfaz unimodal, dentro del constructo de **Utilidad percibida**, se observan diferencias en la valoración del ítem "*El robot me da más control sobre las tareas*", las personas participantes menores de 30 años otorgaron una valoración promedio de 5,40 sobre 7, mientras que aquellos mayores de 30 años asignaron una valoración de 4 sobre 7.

En el constructo de **Facilidad de uso percibida**, se identifican diferencias en la valoración de dos ítems. En el ítem "*El robot es amigable*", las personas participantes menores de 30 años valoraron este aspecto con un promedio de 6,10 sobre 7, mientras que las personas participantes mayores de 30 años otorgaron una valoración de 4,6 sobre 7. De manera similar, en el ítem "*El modo de control del robot me permite recuperarme de los errores rápida y fácilmente*", las personas participantes menores de 30 años asignaron una valoración promedio de 5,70 sobre 7, mientras que aquellos mayores de 30 años valoraron este aspecto con un promedio de 4,9 sobre 7.

Dentro de la categoría de **Controlabilidad**, se destacan diferencias en la valoración de tres ítems. En el ítem "*Puedo configurar el robot según mis necesidades*", las personas participantes menores de 30 años otorgaron una valoración promedio de 4,90 sobre 7, mientras que las personas participantes mayores de 30 años asignaron una valoración de 2,5 sobre 7. En el ítem "*Si el robot está haciendo algo mal, podría encontrar una manera de cambiar su comportamiento*", las personas participantes menores de 30 años asignaron una valoración promedio de 4,10 sobre 7, en contraste con las personas participantes mayores de 30 años, quienes valoraron este aspecto con un promedio de 2,1 sobre 7. De manera similar, en el ítem "*Podría desplegar el robot de una manera específica para ahorrar tiempo*", se observa una diferencia significativa en las valoraciones. Las personas participantes menores de 30 años otorgaron una valoración promedio de 4,70 sobre 7, mientras que aquellos mayores de 30 años asignaron una valoración de 2,7 sobre 7.

En el constructo de **Aprendibilidad**, se destaca una disparidad en la valoración del ítem "*He adquirido destreza con el robot fácilmente*". Las personas participantes menores de 30

años asignaron una valoración promedio de 6,70 sobre 7, mientras que las personas participantes mayores de 30 años otorgaron una calificación de 5,70 sobre 7.

En el constructo de **Actitud e intención de uso**, se identifica una diferencia significativa en el ítem "*Podría conseguir que un robot realizara una tarea específica*". Las personas participantes menores de 30 años otorgaron una valoración promedio de 6,40 sobre 7, mientras que las personas participantes mayores de 30 años asignaron una valoración de 4,1 sobre 7.

En la interfaz multimodal, en la categoría de **Facilidad de uso percibida**, se aprecian diferencias en la valoración de un ítem. En el ítem "*El modo de control del robot me permite recuperarme de los errores rápida y fácilmente*", las personas participantes menores de 30 años valoraron este aspecto con un promedio de 5,70 sobre 7, mientras que las personas participantes mayores de 30 años otorgaron una calificación de 4,1 sobre 7.

Dentro de la categoría de **Controlabilidad**, se identifican diferencias en la valoración de cuatro ítems. En el ítem "*Puedo configurar el robot según mis necesidades*", las personas participantes menores de 30 años asignaron una valoración promedio de 4,5 sobre 7, en contraste con las personas participantes mayores de 30 años, quienes otorgaron una calificación de 2 sobre 7. En el ítem "*Puedo ajustar el robot para que me ayude a realizar mis tareas.*", las personas participantes menores de 30 años valoraron este aspecto con un promedio de 4,60 sobre 7, mientras que las personas participantes mayores de 30 años asignaron una calificación de 2,3 sobre 7. En el ítem "*Puedo resolver problemas con la ayuda del robot*", las personas participantes menores de 30 años otorgaron una calificación promedio de 5,90 sobre 7, en contraste con las personas participantes mayores de 30 años, quienes valoraron este aspecto con un promedio de 3,8 sobre 7. De manera similar, en el ítem "*Podría desplegar el robot de una manera específica para ahorrar tiempo*", se observa una diferencia significativa en las valoraciones. Las personas participantes menores de 30 años asignaron una calificación promedio de 4,9 sobre 7, mientras que las personas participantes mayores de 30 años otorgaron una calificación de 2,6 sobre 7.

Dentro del constructo de **Actitud e intención de uso**, se identifica una diferencia en el ítem "*Podría conseguir que un robot realizara una tarea específica*". Las personas participantes menores de 30 años han otorgado una valoración promedio de 6,2 sobre 7, mientras que las personas participantes mayores de 30 años han asignado una calificación de 3,8 sobre 7.

En el constructo de **Satisfacción**, se observa una disparidad en el ítem "**El robot me hace sentir incómodo/da*". Las personas participantes menores de 30 años valoraron este aspecto con una calificación promedio de 2,7 sobre 7, mientras que las personas participantes mayores de 30 años otorgaron una valoración de 5,5 sobre 7.

Tabla 84: Resultados T-test por ítem (diferencias significativas por género, procedimiento, experiencia y edad) considerando $\alpha = 0,05$. Se resaltan en negrita las diferencias significativas.

Categoría	Ítem	Interfaz Unimodal				Interfaz Multimodal			
		Mujeres / Hombres	Procedi miento I /Procedi miento II	Sin experienci a/Con experienci a	Menores de 30/Mayor es de 30	Mujeres / Hombres	Procedi miento I /Procedi miento II	Sin experiencia/ Con experiencia	Menores de 30/Mayores de 30
Utilidad percibida	El robot mejora mi rendimiento en la tarea.	0,7951	0,7947	0,3380	0,7951	0,8379	0,0160	0,3679	0,2998
	El robot me ayuda a ser más productivo.	0,6162	0,3104	0,5615	1,0000	0,6974	0,0159	0,0436	0,1123
	El robot me ayuda a completar la tarea en menos tiempo.	0,8201	0,8203	0,3530	0,4934	0,3943	0,0819	0,4938	0,1931
	El robot me da más control sobre las tareas.	0,3936	0,1478	0,0472	0,0382	0,3934	0,0189	0,3519	0,8656
	El robot me ayuda a realizar las tareas más fácilmente.	0,6348	0,2635	0,1636	0,0735	0,7035	0,0508	0,2057	0,7035
	El robot satisface mis necesidades.	0,8559	0,5855	0,2241	0,0929	1,0000	0,0889	1,0000	1,0000
Facilidad de uso percibida	Usar el robot es fácil.	0,6071	0,6052	1,0000	0,3007	0,4709	0,1394	1,0000	0,4703
	Usar el robot simple.	0,3138	0,3067	0,8750	0,3077	0,7226	0,0635	0,8271	0,7225
	El robot es amigable.	0,8834	0,6602	0,1919	0,0235	0,5877	0,0042	0,0964	0,8569
	Usar el robot me ayuda.	0,4252	0,7916	0,3399	0,1756	0,8007	0,0690	0,5726	0,8007
	Puedo usar el robot siempre con éxito.	1,0000	0,4400	0,6330	0,2517	0,4030	0,8681	0,9407	0,0585
	Puedo utilizar el robot sin instrucciones escritas.	0,4183	0,4183	0,6884	0,4183	0,4332	0,9114	0,1959	0,7378
	El modo de control del robot requiere el menor número de pasos posible para lograr lo que quiero hacer con él.	0,2513	1,0000	0,6945	0,2552	0,3094	0,3094	0,9291	0,3116
	El modo de control del robot me permite recuperarme de los errores rápida y fácilmente.	0,6604	0,8838	0,5688	0,0177	0,8117	0,4726	0,8990	0,0454
Conseguir un objetivo concreto con la ayuda del robot no será un problema para mí.	0,3585	0,5844	0,6920	0,0900	0,6961	1,0000	0,6269	0,2324	
Seguridad percibida	Me siento seguro/ra cuando utilizo el robot.	0,4655	0,8087	0,6650	0,8087	1,0000	0,0101	0,0088	0,1457
	Durante la tarea, confiaba en que el robot no me haría daño.	1,0000	0,5561	0,2350	0,5561	0,8061	0,2116	0,0758	0,4584

Categoría	Ítem	Interfaz Unimodal				Interfaz Multimodal			
		Mujeres / Hombres	Procedimiento I /Procedimiento II	Sin experiencia/Con experiencia	Menores de 30/Mayores de 30	Mujeres / Hombres	Procedimiento I /Procedimiento II	Sin experiencia/Con experiencia	Menores de 30/Mayores de 30
	*Durante la tarea, me sentía ansioso/sa mientras utilizaba el robot.	0,0889	0,0148	0,4036	0,6826	0,6219	0,0214	0,1472	0,2714
	*Durante la tarea, experimenté incomodidad al utilizar el robot.	0,2409	0,0008	0,1272	0,2409	0,4595	0,0002	0,1862	0,0532
Controlabilidad	El robot es fácil de controlar.	0,5744	0,5744	0,1261	0,5744	0,2758	0,2745	0,7628	0,5166
	Durante la tarea, he sentido que tenía el control del robot.	0,4943	1,0000	0,3978	0,4954	0,2809	0,2786	0,8195	0,5928
	Tengo mucha confianza en mi capacidad para controlar el robot.	0,7955	0,6031	0,5506	0,6035	0,6215	0,2418	0,9325	0,6218
	Puedo configurar el robot según mis necesidades.	0,8402	0,0163	0,0002	0,0077	0,7327	0,1272	0,0002	0,0011
	Puedo ajustar el robot para que me ayude a realizar mis tareas.	0,8419	0,0596	0,0021	0,0596	0,7142	0,0542	0,0000	0,0012
	Puedo resolver problemas con la ayuda del robot.	0,6162	0,3119	0,8889	0,0685	0,6925	0,2290	0,5747	0,0019
	Si el robot está haciendo algo mal, podría encontrar una manera de cambiar su comportamiento.	0,8399	0,0559	0,1608	0,0322	0,6811	0,5361	0,1823	0,0872
Podría desplegar el robot de una manera específica para ahorrar tiempo.	1,0000	0,1263	0,1288	0,0229	0,9189	0,2540	0,1187	0,0112	
Aprendibilidad	Puedo aprender fácilmente a utilizar el robot.	0,2290	0,4763	0,9183	0,8136	0,8136	0,4763	0,5935	0,8137
	Puedo recordar fácilmente cómo se utiliza el robot.	0,3307	0,7486	0,8826	0,3307	0,5674	0,0741	1,0000	0,5676
	He adquirido destreza con el robot fácilmente.	0,4133	0,6855	0,2303	0,0369	0,2214	0,0109	0,1886	0,4153
Actitud e intención de uso	Podría utilizar el robot en mi trabajo diariamente.	0,4878	0,0311	0,1153	0,8180	0,6331	0,1441	0,4648	1,0000
	Podría conseguir que un robot realizara una tarea específica.	0,1262	0,3719	0,1703	0,0202	0,4999	0,1089	0,1067	0,0036
	Si tengo la opción de elegir, prefiero trabajar con el robot.	0,1141	0,4407	0,6519	0,4413	0,7127	0,0543	0,8251	1,0000

Categoría	Ítem	Interfaz Unimodal				Interfaz Multimodal			
		Mujeres / Hombres	Procedi miento I /Procedi miento II	Sin experienci a/Con experienci a	Menores de 30/Mayor es de 30	Mujeres / Hombres	Procedi miento I /Procedi miento II	Sin experiencia/ Con experiencia	Menores de 30/Mayores de 30
	Creo que utilizaría el robot en futuras tareas.	0,4600	0,2640	0,7636	0,7135	0,8414	0,0619	0,6973	0,8413
	Durante la tarea, confiaba en que el robot hiciera lo correcto en el momento adecuado.	0,4314	0,7947	0,2272	0,0598	0,1576	0,0619	0,6973	0,5478
Satisfacción	Estoy satisfecho/cha con el robot.	1,0000	1,0000	0,4831	0,3468	0,4332	0,0104	0,6584	0,4370
	Durante la tarea, el robot y yo trabajamos bien juntos.	0,7008	0,7027	0,4660	0,1180	0,3013	0,2964	0,8727	0,7313
	El robot funciona como yo quiero que funcione.	0,3817	0,8619	0,0700	0,0507	0,2141	1,0000	0,3446	0,0588
	El uso del robot es armonioso.	0,5069	0,7415	0,4006	0,0870	1,0000	0,0373	0,8422	0,7782
	* El robot me hace sentir incómodo/da.	0,7731	0,0033	0,3768	0,0901	0,7129	0,7659	0,0001	0,0040
	Me gusta usar el robot.	0,1625	0,3229	0,9376	0,5583	0,5561	0,0053	0,2592	0,8453
	Me siento cómodo/da con el robot.	0,3717	0,3728	0,8678	0,6597	0,4927	0,0056	0,3530	0,8201

6.3.4 Discusión

En esta sección se procede a interpretar y analizar en detalle los resultados obtenidos del estudio, enfocándose en el impacto de las interfaces unimodales y multimodales en los aspectos relacionados con los factores humanos.

6.3.4.1 *Perspectiva de Utilidad y Experiencia*

El análisis de los resultados obtenidos en el estudio revela que, en el procedimiento I, la interfaz multimodal obtuvo valoraciones más altas en utilidad percibida, facilidad de uso percibida, seguridad percibida, aprendibilidad y satisfacción en comparación con la interfaz unimodal. Sin embargo, en el procedimiento II, las diferencias entre las interfaces fueron menos pronunciadas. Aunque la interfaz unimodal obtuvo valoraciones ligeramente más elevadas en utilidad percibida y seguridad percibida, no se encontraron diferencias significativas. Esto podría deberse a varios factores que podrían haber influido en los resultados.

Se postula que en el procedimiento II, las personas participantes ya habían experimentado la interfaz multimodal. Al haber tenido previa exposición a la interfaz multimodal, se deduce que las personas participantes podrían haber adquirido cierta familiaridad y conocimiento sobre cómo interactuar con el sistema. Esto podría haberles permitido adaptarse más rápidamente a la interfaz unimodal en el procedimiento II, lo que a su vez podría haber influido en sus valoraciones.

Además, es importante tener en cuenta que la tarea y el sistema en general tenían un alto grado de aprendibilidad. Es decir, las personas participantes tenían la oportunidad de aprender y familiarizarse con el sistema a lo largo del estudio, independientemente de la interfaz utilizada. Esto significa que incluso si inicialmente experimentaron diferencias en la percepción entre las interfaces, es posible que estas diferencias se hayan atenuado a medida que las personas participantes se familiarizaban y se adaptaban al sistema en general.

La preferencia por la interfaz multimodal en términos de utilidad percibida sugiere que las personas participantes encontraron beneficios adicionales al interactuar con una interfaz que combina dos tipos de interfaces, visual y voz, en comparación con la interfaz unimodal, que utilizaba solo la interfaz visual. Esto podría deberse a que la interfaz multimodal proporciona una mayor flexibilidad y variedad en las formas de interactuar, lo que a su vez puede mejorar la eficacia y eficiencia de la interacción.

En cuanto a la secuencia en la ejecución de las tareas, el hecho de que las personas participantes realizaran la tarea inicialmente con la interfaz multimodal y posteriormente con la unimodal puede haber influido en los resultados. Es posible que la familiaridad adquirida durante la primera ejecución con la interfaz multimodal haya mejorado la eficiencia y adaptación en la segunda ejecución con la interfaz unimodal. Esto puede explicar por qué no se encontraron diferencias significativas en las dimensiones evaluadas entre las interfaces. El alto grado de aprendibilidad del sistema también puede haber contribuido a la agilidad en la ejecución de la tarea, independientemente de la interfaz utilizada en la segunda instancia.

Sin embargo, es importante tener en cuenta que estos resultados son específicos para la muestra y el contexto del estudio en cuestión. En este sentido, se recomienda realizar investigaciones adicionales con muestras más amplias y en diferentes entornos para obtener una comprensión más completa de los efectos de las interfaces unimodales y multimodales en la interacción y evaluación de las personas usuarias.

6.3.4.2 *Análisis de las diferencias entre grupos*

En esta sección, se presentan los resultados de un análisis de diferencias entre condiciones en las tareas. Se han considerado factores como el procedimiento empleado (procedimiento I o procedimiento II), el género de las personas participantes (mujeres o hombres), su edad (menores de 30 o mayores de 30), o su experiencia con robots (sin experiencia o con experiencia) para identificar posibles diferencias significativas en los indicadores recopilados. Los análisis se realizaron utilizando pruebas t de Student, con un nivel de significancia $\alpha = 0,05$. Los resultados se muestran en la Tabla 85.

En el índice de producción no se observa diferencia significativa en ninguna de las interfaces en ninguna de las condiciones.

En cuanto a las respuestas fisiológicas, tan solo se observa una diferencia significativa en el *engagement* de la interfaz unimodal entre procedimientos, siendo las personas participantes del procedimiento I los que han tenido valores de *engagement* mayores que los del procedimiento II en la interfaz unimodal.

En cuanto a la respuesta perceptual se observa diferencia significativa en los siguientes aspectos:

- La **utilidad percibida** en la interfaz multimodal entre procedimientos, siendo las personas participantes del procedimiento I los que lo han valorado muy por encima. Esta diferencia sugiere que las personas participantes del procedimiento I al comparar la interfaz multimodal con la unimodal, indican cómo la perciben bastante más útil.
- La **seguridad percibida** muestra diferencias significativas tanto en la interfaz unimodal como en la multimodal entre procedimientos. En este sentido, se observa que la interfaz unimodal es percibida como más segura para las personas participantes del procedimiento I que para las personas participantes del procedimiento II. En la misma línea, las personas participantes del procedimiento I valoran más alto la seguridad percibida en la interfaz multimodal que las personas participantes del procedimiento II. Esto puede deberse a la propia comparación que realizan. Las personas participantes del procedimiento I valoran mucho mejor la interfaz multimodal al compararla con la unimodal. Los del procedimiento II, valoran mucho peor la interfaz unimodal al compararla con su previa interacción con la interfaz multimodal. Por otro lado, también se observa diferencia significativa en la seguridad percibida de la interfaz multimodal entre las personas con y sin experiencia, siendo las personas sin experiencia las que la perciben como más segura.
- Se observan diferencias significativas en la **controlabilidad** de la interfaz unimodal y multimodal entre edades, siendo en ambos casos las personas participantes menores de 30 los que han sentido mayor controlabilidad. También se observa diferencia significativa en la controlabilidad de la interfaz unimodal según la experiencia, siendo las personas participantes con experiencia los que lo han valorado con mayor puntuación.
- En cuanto a la **aprendibilidad** solo se observa diferencia significativa en la interfaz multimodal entre procedimientos, siendo las personas participantes del procedimiento I los que han otorgado mayores valoraciones.
- Finalmente, se observa una diferencia significativa en la **satisfacción** de la interfaz multimodal entre procedimientos, siendo las personas participantes del procedimiento I los que han otorgado nuevamente mayores valoraciones.

Tabla 85: Resultados T-test por indicador recogido (diferencias significativas por procedimiento, género, edad y experiencia) considerando $\alpha = 0,05$. Se resaltan en negra las diferencias significativas.

		Unimodal				Multimodal			
		Procedimiento	Género	Edad	Experiencia	Procedimiento	Género	Edad	Experiencia
Desempeño	Índice de producción	0,2443	0,2088	0,0935	0,4271	0,0817	0,3255	0,8550	0,1266
	Engagement	0,0376	0,7591	0,6632	0,0807	0,3462	0,4694	0,2233	0,7419
Respuestas fisiológicas	Memorización	0,9267	0,1007	0,4933	0,4068	0,2432	0,3362	0,1211	0,7754
	Valencia	0,0903	0,6208	0,7965	0,1891	0,3011	0,4289	0,1140	0,5282
	Carga mental	0,5275	0,8627	0,2823	0,2316	0,7014	0,5962	0,2899	0,9182
Respuestas perceptuales	Utilidad percibida	0,6710	0,7343	0,1678	0,0903	0,0089	0,6353	0,3816	0,1937
	Facilidad de uso percibida	0,6507	0,4488	0,0651	0,3849	0,1313	0,9493	0,2203	0,9463
	Seguridad percibida	0,0035	0,1244	0,5310	0,2018	0,000002	0,8568	0,0577	0,0095
	Controlabilidad	0,1119	0,9669	0,0226	0,0331	0,3999	0,6238	0,0019	0,0877
	Aprendibilidad	0,8436	0,2268	0,4287	0,5645	0,0490	0,5327	0,5311	0,8827
	Actitud e intención de uso	0,6349	0,1445	0,1235	0,4084	0,3111	0,4815	0,1832	0,5319
	Satisfacción	0,1532	0,3642	0,4513	0,6739	0,0007	0,6233	0,7692	0,3412

6.3.4.3 El impacto de la aprendibilidad de la tarea

Los indicadores de desempeño ofrecen una visión objetiva sobre la dinámica de la interacción entre el individuo y el robot en un entorno colaborativo. A través de los datos obtenidos, es evidente que la influencia más marcada recae en la capacidad de las personas participantes para aprender la tarea, más que en el tipo específico de interfaz empleado. De hecho, se puede inferir que, independientemente de la interfaz utilizada, la mayoría de las personas participantes logró un desempeño superior en la segunda ejecución de la tarea.

Este fenómeno sugiere que la experiencia previa y la familiarización con la tarea tienen un efecto positivo en el rendimiento global. La mejora en los resultados en la segunda ejecución podría atribuirse a la adquisición de habilidades y estrategias más efectivas, lo cual resalta la importancia de la adaptación y el aprendizaje continuo en la HRI.

Es fundamental destacar que la relación entre el tipo de interfaz y el desempeño no es necesariamente directa, y que otros factores, como la experiencia previa del participante y la familiaridad con las tecnologías de interfaz, podrían influir significativamente en los resultados observados.

6.3.4.4 Reducción de memorización y carga mental con la interfaz multimodal

Dentro del análisis de los indicadores fisiológicos, resulta notable la influencia de las interfaces unimodales y multimodales en los niveles de memorización y carga mental experimentados por las personas participantes. Una observación crucial es que esta influencia se manifiesta de manera constante, independientemente de la secuencia de uso de las interfaces, lo cual añade una capa adicional de relevancia a los resultados.

La disminución en los niveles de memorización y carga mental en presencia de la interfaz multimodal podría indicar una mayor eficiencia cognitiva durante la interacción. La capacidad de intercambiar información y recibir retroalimentación a través de múltiples modalidades puede haber facilitado la asimilación de los procedimientos y detalles de la tarea.

Esto puede ser especialmente significativo en un entorno industrial, donde la precisión y la rapidez son esenciales.

El descenso en la carga mental también puede ser interpretado como una señal de una experiencia de interacción menos demandante. La interfaz multimodal, al proporcionar múltiples canales de información, podría haber reducido la carga cognitiva total al distribuir la información de manera más efectiva. Como resultado, las personas participantes podrían haber experimentado una sensación de fluidez y comodidad durante la ejecución de la tarea.

6.3.4.5 La perspectiva de género

Una de las consideraciones esenciales en la evaluación de la HRI en entornos industriales es analizar las posibles diferencias de género en la percepción y el desempeño. Es importante analizar los datos desde esta perspectiva para hacer frente a la brecha digital de género ya existente.

Al explorar los indicadores de desempeño en relación con el género de las personas participantes, los resultados muestran una notable ausencia de diferencias significativas. Tanto hombres como mujeres demostraron respuestas similares en términos de mejora de rendimiento en la segunda ejecución de la tarea, independientemente del tipo de interfaz utilizado. Esta observación destaca que el género no parece ser un factor determinante en la adaptación y el aprendizaje de la tarea colaborativa en este contexto industrial específico.

En la evaluación de los indicadores fisiológicos, se revela que las diferencias entre las interfaces unimodales y multimodales en términos de memorización y carga mental no parecen verse afectadas por el género. Ambos géneros experimentan una reducción consistente en los niveles de memorización y carga mental con la interfaz multimodal, independientemente de la secuencia de uso.

Al analizar los resultados del cuestionario HUROY desagregados por género, se observa que no se encontraron diferencias significativas entre hombres y mujeres en ninguna de las interfaces evaluadas en los dos procedimientos. Ambos géneros mostraron respuestas similares en cuanto a utilidad percibida, facilidad de uso percibida, seguridad percibida, controlabilidad, aprendibilidad, actitud e intención de uso, y satisfacción. Esto indica que la percepción de las interfaces no se vio influenciada por el género de las personas participantes.

Los hallazgos en relación con el género en este estudio sugieren que, al menos en este contexto particular, las interfaces unimodales y multimodales presentan una adaptación y aceptación similares entre hombres y mujeres.

Principales conclusiones del capítulo

- Se valida el modelo ITPX-ROBOTS mediante la realización de dos casos de estudio. En el primer caso de estudio, además de validar el modelo, se investiga cómo las variables cinemáticas, específicamente la velocidad y la aceleración, afectan a los factores humanos. En el segundo caso de estudio, además de validar el modelo, se examina el impacto de utilizar una interfaz multimodal en comparación con una unimodal en la HRI industrial.
- Los resultados de la validación indican que el modelo ITPX-ROBOTS es un enfoque adecuado para evaluar la HRI en entornos industriales y proporciona una evaluación precisa de los aspectos relevantes en esta área.

Caso de estudio 1: Influencia de las variables cinemáticas

- El cuestionario UEQ-S no se muestra como la herramienta más idónea para obtener una comprensión más profunda de las percepciones. Presenta algunas limitaciones, como la falta de capacidad para captar toda la experiencia del usuario, en contraste con su versión original, el UEQ. Además, se observa que las preguntas no reflejan de manera precisa la interacción entre las personas y los robots en la vida real. Esto conduce a la necesidad de identificar un cuestionario más apropiado para su uso en entornos industriales con robots. Dado que no se encontró una alternativa adecuada en la literatura existente, se optó por desarrollar y validar el cuestionario HUOX.
- Los hallazgos señalan que, en general, el aumento de las velocidades y aceleraciones tiende a generar una mayor carga de trabajo mental y a una menor valencia emocional y engagement.
- La extensión del tiempo de interacción con el robot puede influir en las reacciones emocionales, ya que interacciones de mayor duración suelen asociarse con calificaciones más bajas en la activación y el impacto.
- Las personas participantes otorgaron calificaciones más altas en cuanto a la calidad práctica, la calidad hedónica y la utilidad percibida cuando el robot se movía a una mayor velocidad. Sin embargo, al mismo tiempo, esta mayor velocidad también tuvo un impacto negativo en las calificaciones relacionadas con la fiabilidad y la capacidad de control del robot.
- La velocidad y la aceleración son variables críticas a tener en cuenta para optimizar la UX en entornos de trabajo. Los resultados de los experimentos muestran que una velocidad excesivamente alta tiene un impacto negativo en la experiencia de interactuar con el robot, al igual que una velocidad excesivamente baja. Como resultado, se ha definido el concepto de la "campana de velocidad de HRI", que se ilustra en la Figura 57. Esta campana de velocidad muestra que una velocidad y aceleración del robot demasiado altas se perciben como inseguras y poco confiables por parte de la persona, mientras que una velocidad y aceleración demasiado bajas también generan desconfianza en la persona.
- El orden en que se llevan a cabo las tareas tiene un impacto en las expectativas y respuestas emocionales de las personas que participan en la interacción con el robot. Esto se evidencia en las diferencias en las puntuaciones entre los procedimientos LR y RL, donde el orden de las tareas parece influir en los resultados. Es crucial tener en cuenta la influencia de la facilidad de aprendizaje al diseñar HRI, ya que esta puede afectar tanto la UX como el desempeño. Al comprender el efecto del orden de las tareas en las interacciones entre personas y robots, es posible mejorar el diseño de los robots para satisfacer de manera más efectiva las necesidades y expectativas de las personas, aumentando así su bienestar en el entorno industrial moderno. El estudio ha identificado una "paradoja de la aprendibilidad" en la HRI (Figura 58), donde los tiempos de interacción más largos mejoran el rendimiento y la facilidad de aprendizaje de los participantes, pero al mismo tiempo reducen la calidad de la UX.
- Se han observado diferencias entre géneros tan solo en el número de

errores en la TL, siendo las mujeres las que han cometido menor número de errores. Sin embargo, debido al tamaño de la muestra y la posible influencia de otros factores, se necesita más investigación para comprender y abordar completamente esta diferencia.

- Según los indicadores de rendimiento, se observa que cuanto más rápido se movía el robot, más rápido se completaba la tarea. Por lo tanto, en todos los casos, los participantes fueron más rápidos en la TR. Sin embargo, el orden en que se realizan las tareas influye en los resultados en un estudio de rendimiento de tareas. Las personas que llevaron a cabo las tareas en orden inverso tardaron más en completar la tarea y cometieron más errores en la primera tarea que en la segunda. La segunda tarea se completó de manera más efectiva, con menos errores. Esto sugiere que la facilidad de aprendizaje de la interacción mejora el rendimiento.

Caso de estudio 2: Impacto de las interfaces unimodales vs. multimodales

- Se emplea la versión avanzada de HEUROBOX para realizar la evaluación por parte de las personas expertas, y se obtiene una puntuación total del 61,05% para el sistema con interfaz unimodal y del 61,6% para el sistema con interfaz multimodal. Dado que la única diferencia entre ambos sistemas reside en la interfaz misma, las variaciones en las puntuaciones se limitan principalmente a la categoría de interfaces. Esta diferencia es mínima, ya que la categoría de interfaces tiene un peso relativamente bajo en el puntaje total según el indicador HEUROBOX.
- En cuanto a los **indicadores de desempeño**, se destaca el impacto de la aprendibilidad de la tarea, ya que se observa una mejora en los indicadores de desempeño en la tarea que se ejecuta en segundo lugar en ambos procedimientos. Esto sugiere que la familiaridad previa con una modalidad de interfaz específica puede mejorar el rendimiento en la segunda ejecución de la tarea, ya que las personas participantes pueden desarrollar estrategias más eficientes y adaptarse mejor.
- En cuanto al **engagement**, se observa que las personas participantes experimentaron un mayor nivel de involucramiento emocional y atención al interactuar con la interfaz unimodal en el Procedimiento I. Sin embargo, en el Procedimiento II, se invierte la tendencia, y las personas participantes demuestran un mayor nivel de *engagement* al interactuar con la interfaz multimodal. Esto indica que el *engagement* parece ser más alto al inicio de la interacción con el robot, pero tiende a disminuir con el tiempo.
- En cuanto a la **memorización**, se observa que en ambos procedimientos las personas participantes experimentan mayor nivel de memorización al interactuar con la interfaz unimodal. Se puede concluir, por lo tanto, que la interfaz multimodal, independientemente del procedimiento está relacionada con una disminución de los niveles de memorización en comparación con una interfaz unimodal.
- En lo que respecta a la **valencia emocional**, se observa que las personas participantes experimentan emociones más negativas al utilizar la interfaz unimodal en comparación con la interfaz multimodal. Estos hallazgos sugieren que la interfaz multimodal proporciona una experiencia emocional más positiva en general.
- En lo que concierne a la **carga mental**, se observa que las personas

participantes experimentan una carga mental significativamente más elevada al interactuar con la interfaz unimodal en comparación con la interfaz multimodal. Estos resultados sugieren que la interfaz multimodal tiende a reducir la carga mental percibida por las personas usuarias, lo que podría indicar una experiencia de interacción más eficiente y menos demandante en términos cognitivos en comparación con la interfaz unimodal.

- En lo que concierne a los **indicadores perceptuales**, es importante destacar que todos los aspectos evaluados en el cuestionario HUOX muestran calificaciones notablemente más positivas cuando las personas interactúan con la interfaz en la segunda posición. En otras palabras, en el Procedimiento I, la interfaz multimodal recibe valoraciones favorables en la mayoría de los ítems, mientras que en el Procedimiento II, es la interfaz unimodal la que obtiene calificaciones más altas. Esta tendencia constante sugiere que, al realizar la tarea por segunda vez, las personas se sienten más cómodas y familiarizadas con la interfaz, lo que puede haber influido positivamente en sus percepciones y evaluaciones, independientemente de las características específicas de cada interfaz. Es relevante mencionar que la facilidad de aprendizaje de la tarea también se refleja en los indicadores de rendimiento, lo que respalda aún más la idea de que la familiaridad y la comodidad desempeñan un papel crucial en la mejora de las calificaciones de las personas en la segunda ejecución de la tarea.
- No se han identificado diferencias estadísticamente significativas entre géneros en ninguno de los aspectos evaluados. Los resultados de este estudio sugieren que, al menos en este contexto específico, tanto las interfaces unimodales como las multimodales son igualmente aceptadas y adoptadas tanto por hombres como por mujeres. Esto plantea la posibilidad de avanzar hacia la reducción de la brecha digital de género en entornos industriales de HRI.

Capítulo VII

Conclusiones y líneas futuras

7 CONCLUSIONES Y LÍNEAS FUTURAS

Este capítulo se estructura en cinco secciones. En primer lugar, se presentan las conclusiones más destacadas que emergen a lo largo de esta investigación, proporcionando una visión global del trabajo. En segundo lugar, se lleva a cabo una revisión de la validación de las hipótesis previamente definidas. A continuación, se enumeran las contribuciones más significativas que han surgido durante el desarrollo de esta tesis. En el cuarto apartado, se realiza una reflexión crítica sobre las limitaciones que han acompañado a la investigación. Por último, se esbozan las líneas futuras identificadas.

7.1 Conclusiones generales

A continuación, se recogen las conclusiones principales de la investigación, recogidas en base a los capítulos que componen el trabajo.

Las principales conclusiones extraídas del Capítulo I se resumen de la siguiente manera:

- La industria manufacturera está experimentando una transformación tecnológica en las formas de Industria 4.0 y 5.0. La primera busca eficiencia y automatización, mientras que la segunda se enfoca en la sostenibilidad y el bienestar de las personas. La transición hacia la Industria 5.0 plantea desafíos en la definición de los roles de las personas en entornos tecnológicos en evolución. Se destaca la necesidad de métodos de evaluación específicos para garantizar la seguridad y la eficacia en entornos industriales reales.
- En la nueva industria, los robots desempeñan un papel esencial debido a su relevancia en la implementación. Sin embargo, las personas serán cruciales para garantizar la flexibilidad y la adaptación a las cambiantes necesidades del mercado y la personalización de productos. En este sentido, la HRI promueve una relación simbiótica en las fábricas inteligentes, donde los robots asumen tareas repetitivas y peligrosas, mientras que las personas se encargan del pensamiento crítico. A medida que los robots se integran en entornos de trabajo compartidos, es fundamental garantizar que los robots sean fáciles de usar y que las personas operarias respondan positivamente a la tecnología. Para lograr una colaboración exitosa entre personas y robots, se deben cumplir dos criterios: demostrar que los diseños centrados en las personas son más eficaces y eficientes, y garantizar que la respuesta de las personas operarias a la tecnología sea positiva.
- El TAM (Davis, 1989, 1993) es un marco teórico ampliamente utilizado para comprender y predecir la aceptación de la tecnología por parte de las personas usuarias. El modelo HRCAM (Bröhl et al., 2019), se basa en el modelo TAM y se centra específicamente en la aceptación de la colaboración persona-robot. Reconoce que los robots son percibidos y evaluados de manera diferente en comparación con otras tecnologías debido a sus características únicas, como el movimiento autónomo y la interacción con el entorno. El HRCAM identifica diversos predictores de la utilidad percibida y la facilidad de uso percibida en el contexto de la HRI, incluyendo la relevancia laboral, la norma subjetiva, la calidad de salida, la demostrabilidad, el control externo, la autoeficacia, la ansiedad ante los robots, el disfrute percibido, la seguridad percibida y más.
- Coronado et al. (2022) proponen un modelo conceptual holístico específicamente diseñado para la HRI. Este modelo clasifica los factores en

usabilidad, UX, hedonomía y ergonomía, y muestra las relaciones y límites entre estos conceptos. Además, agrupan los factores hedonómicos en dos grupos: aquellos influenciados por aspectos emocionales y aquellos donde tanto los aspectos emocionales como los cognitivos son relevantes.

Las principales conclusiones extraídas del Capítulo II se resumen de la siguiente manera:

- La SLR revela la existencia de cinco marcos teóricos recientemente creados. La aparición de estos marcos teóricos refleja un creciente interés en el área de estudio. El marco teórico propuesto por Cohen et al (2018) subraya la importancia del *feedback* adecuado en la HRI. La selección del modo de reacción basado en la observación y el análisis de elementos relevantes es esencial para mejorar la UX. Sin embargo, queda abierta la oportunidad de investigar cómo personalizar este *feedback*, considerando las acciones del robot y las emociones de la persona usuaria.
- Villani et al. (2018) enfocan su marco teórico en la seguridad, destacando la necesidad de interfaces intuitivas y la importancia de la robótica afectiva para aliviar la carga cognitiva de la persona usuaria. Por otro lado, la ergonomía ha surgido como un tema crítico en la HRI, con especial énfasis en la ergonomía física y la necesidad de abordar la ergonomía cognitiva de manera más completa.
- El marco teórico de Meissner et al. (2020) identifica los factores influyentes en la aceptación de las personas trabajadoras al interactuar con robots. La percepción de riesgo, los beneficios percibidos y las emociones son factores clave que influyen en la actitud de las personas hacia los sistemas. Los marcos teóricos identificados reconocen la importancia de la UX y la aceptación tecnológica en la HRI. De hecho, Lindblom & Alenljung (2020) proponen el marco ANEMONE, que se centra en la comprensión de la percepción y la acción del robot por parte de la persona usuaria. Esto destaca la necesidad de investigar cómo las personas usuarias perciben, anticipan y evalúan las acciones del robot y cómo se puede mejorar esta percepción.
- Coronado et al. (2022) señalan la importancia de evaluar los aspectos hedonómicos de la HRI en la industria mediante el desarrollo de medidas y métricas específicas que puedan evaluar las emociones positivas y negativas de las personas usuarias.
- A pesar de los avances en la creación de los marcos teóricos, se identifica la falta de un método de evaluación estructurado que se adapte a diferentes contextos y etapas del proceso de diseño en la HRI. El marco propuesto por Prati et al. (2021) se acerca a abordar esta necesidad, pero muestra carencias en la fase de evaluación.
- La calidad de la HRI depende de interfaces intuitivas y eficientes que permitan una comunicación clara entre las personas usuarias y los robots. Esto es especialmente importante en colaboraciones cercanas, donde la facilidad de interacción es esencial. En este sentido, la efectividad de la HRI se basa en definir las interacciones deseadas y el propósito de los intercambios de información, adaptándolos a diferentes contextos. Las interfaces pueden involucrar comunicaciones visuales, gestuales, basadas en voz y lenguaje natural, así como interacciones físicas y hápticas, categorizadas según la naturaleza de la comunicación y las modalidades utilizadas.
- La HRI se ha clasificado en áreas como coexistencia persona-robot, cooperación persona-robot y colaboración persona-robot (HRC), dependiendo

del nivel de interacción. A medida que aumenta el nivel de interacción, también aumenta el riesgo asociado, especialmente en colaboraciones cercanas donde la seguridad es crucial.

- La multimodalidad natural en la HRI introduce un nivel adicional de complejidad en la comunicación, y su efectividad se evalúa en función de su capacidad para facilitar interacciones genuinamente naturales. La evaluación de si se logra este objetivo representa un desafío crítico para la investigación futura en este campo en constante desarrollo.
- En la SLR, no se ha encontrado un modelo validado para la evaluación de la UX y la aceptación tecnológica en la HRI en entornos industriales, lo que subraya la necesidad imperante de llevar a cabo investigaciones en este ámbito específico.
- Se identifica la necesidad de adoptar un enfoque holístico que recoja las percepciones de las personas en todas las fases de la interacción, integre diferentes métodos de medición (cuantitativos y cualitativos), y considere aspectos de género para comprender mejor la UX en HRI.
- Los cuestionarios son las herramientas más utilizadas para la evaluación subjetiva de la HRI, aunque se identifica la carencia de un consenso validado o cuestionario específico para este contexto. Esta falta de instrumentos específicos para evaluar la calidad de la UX en la HRI industrial resalta la oportunidad de desarrollar o adaptar cuestionarios existentes que aborden dimensiones particulares de esta área en constante evolución.
- Se evidencia una tendencia en alza en la incorporación de medidas fisiológicas en la investigación de la HRI, particularmente en estudios recientes. Las mediciones se enfocan primordialmente en el ECG y la EDA, y la tecnología de wearables está desempeñando un papel facilitador en esta integración. El empleo de tecnología psicofisiológica, como el EEG, el ECG y el EDA, emerge como una herramienta valiosa para capturar las respuestas emocionales y cognitivas de las personas durante la interacción con robots. Aunque se observa un aumento en el interés por la inclusión de tecnología fisiológica en la evaluación de la UX, existen limitaciones que deben abordarse, como los costos asociados, la intrusividad y la necesidad de conocimientos especializados.
- Se discute la falta de consenso sobre el tamaño de muestra adecuado en estudios que utilizan monitorización fisiológica en la UX. Se destaca la importancia de investigar y definir tamaños de muestra apropiados.
- La variable de género ha sido subrepresentada en los estudios analizados, y se señala la importancia de incorporar la perspectiva de género en la investigación de HRI para abordar la brecha digital de género y garantizar la equidad en la interacción con robots.

En conjunto, estas conclusiones resaltan la necesidad de una investigación más integral y diversificada en el campo de la HRI en entornos industriales, que tenga en cuenta tanto aspectos objetivos como subjetivos, así como la perspectiva de género, para comprender mejor y mejorar la HRI en estos contextos. Se subraya la necesidad de una evaluación exhaustiva de la HRI en entornos industriales, utilizando una variedad de enfoques, desde tecnología psicofisiológica hasta cuestionarios de UX, para comprender mejor la experiencia de las personas y mejorar la colaboración entre personas y robots en estos entornos. También destaca la necesidad de abordar las limitaciones y desafíos asociados con estas técnicas de evaluación.

Las principales conclusiones extraídas del Capítulo III se resumen de la siguiente manera:

- Se propone un nuevo marco conceptual llamado Human Robot Interaction eXperience (HRIX), un modelo basado en el modelo CUE desarrollado por Thüring & Mahlke (2007). El marco sostiene que la interacción entre una persona y un robot se ve influenciada por tres conjuntos: (i) las características de la persona, (ii) las características del robot, y (iii) las características del contexto. Estos factores juntos determinan el ajuste y la calidad de la interacción, el HRI FIT. Este ajuste entre las características tiene un impacto en los componentes de la UX, que en este marco se definen como la utilidad, la facilidad de uso, el confort, la aprendibilidad y la controlabilidad. Todo ello, tiene unas consecuencias en los factores humanos como la actitud e intención de uso, la confianza, la seguridad percibida y la satisfacción.
- Las características que componen la taxonomía HUBOXT sugieren que la aceptación de la HRI por parte de las personas trabajadoras está influenciada por una compleja interacción de diferentes factores. De acuerdo con Meissner et al. (2020), un sistema robótico seguro y bien diseñado es una condición necesaria pero no suficiente para la aceptación de la HRI en entornos industriales. Tanto los investigadores como las empresas deben considerar la aceptación de la HRC no solo como un problema tecnológico sino también cultural. Para empezar, **las características de la persona** que ejercen influencia en la aceptación de la HRI son (i) las características sociodemográficas, (ii) la motivación, (iii) la autoevaluación, (iv) las experiencias previas y el conocimiento, (v) las normas sociales subjetivas, (vi) la posición en el equipo, (vii) la personalidad y los valores morales y (viii) la intención de comportamiento. Por otro lado, **las características del robot** que ejercen influencia en la aceptación de la HRI son: (i) los atributos de seguridad, (ii) la tarea, (iii) la programación, (iv) el nivel de procesamiento y la memoria, (v) la precisión y repetibilidad, (vi) la carga útil y el alcance, (vii) la velocidad y la aceleración, (viii) la autonomía y el control, y, (ix) la flexibilidad y adaptabilidad. Finalmente, **las características del contexto** que ejercen influencia en la aceptación de la HRI se dividen en las características del micro contexto y las características del macro contexto. Las características del micro contexto son: (i) el diseño físico del espacio de trabajo, (ii) la disposición de la tecnología, (iii) la carga de trabajo, y (iv) el proceso de implementación. Las características del macro contexto son: (i) la cultura organizacional, (ii) el liderazgo y la gestión, (iii) la relación con ejecutivos, y (iv) la relación con los compañeros y las compañeras.
- Se presenta la concepción del modelo de evaluación ITPX-ROBOTS, que se basa en un enfoque holístico de evaluación de la UX en HRI en entornos industriales. Se compone de tres fases: (i) la evaluación previa, (ii) la evaluación durante la interacción y (iii) la evaluación posterior a la interacción. La **evaluación previa** implica la evaluación experta, donde se busca determinar la compatibilidad entre las características de la persona, el robot y el entorno. Se utiliza una serie de heurísticas llamada HEUROBOX para este propósito, que abordan aspectos de seguridad, ergonomía, funcionalidad e interfaces. La **evaluación durante la interacción** se centra en la realización de un test de usabilidad que recopila indicadores de rendimiento y valores fisiológicos. Los indicadores de desempeño incluyen el tiempo de ejecución de la tarea, errores y el índice de producción. Además, se utilizan medidas fisiológicas, como la activación, el impacto, la valencia, la memorización, la atención y el

engagement, para comprender las respuestas emocionales y cognitivas de las personas durante la interacción. Finalmente, en **la evaluación posterior a la interacción**, se recopila la percepción de las personas usuarias mediante el cuestionario HUROY. Este cuestionario aborda constructos como la utilidad percibida, facilidad de uso percibida, seguridad percibida, controlabilidad, aprendibilidad, actitud y satisfacción, proporcionando una visión completa de la experiencia de la persona usuaria.

El Capítulo IV introduce HEUROBOX, una herramienta de evaluación basada en heurísticas desarrollada y validada para la HRI en entornos industriales y que se basa en cuatro categorías diferentes: (i) la seguridad, (ii) la funcionalidad, (iii) la ergonomía y (iv) las interfaces. A continuación, se resumen las principales conclusiones extraídas de este capítulo:

- HEUROBOX es una herramienta que facilita la evaluación de la HRI y permite la identificación de aspectos para mejorar en el sistema en términos de UX. Sin embargo, para obtener una evaluación completa y real de la UX, es apropiado complementarla con otras técnicas de evaluación de UX, como pruebas de usabilidad.
- HEUROBOX aborda la **seguridad** mediante consideraciones de diseño destinadas a prevenir daños tanto a las personas como a los robots. Esto incluye medidas de seguridad como paradas de emergencia, barreras de seguridad y sistemas de detección de colisiones. Además, el diseño del robot debe considerar las limitaciones físicas y cognitivas de las personas usuarias y ser capaz de adaptarse a situaciones inesperadas.
- HEUROBOX, en términos de **ergonomía**, permite evaluar si las personas usuarias puedan operar el robot sin sufrir tensiones físicas indebidas o incomodidades.
- HEUROBOX permite evaluar si la **funcionalidad** del robot permite adaptarse al contexto específico de uso y ser capaz de manejar cualquier situación inesperada que pueda surgir. El robot también debe ser capaz de realizar tareas de manera autónoma, con una intervención humana mínima si es necesario.
- HEUROBOX abarca una serie de consideraciones de diseño en las **interfaces**, incluidas las interfaces visuales, de voz, de gestos y hápticas. Las interfaces visuales deben diseñarse para proporcionar la información necesaria de manera clara y estructurada, con gráficos e iconos simples. Las interfaces de voz deben diseñarse para ser naturales y similares a las humanas, con un diálogo estructurado entre la persona usuaria y el sistema. Las interfaces de gestos deben ser intuitivas y naturales para que la persona usuaria las utilice, mientras que las interfaces hápticas deben proporcionar suficiente retroalimentación táctil al usuario.
- Si bien las cuatro categorías que componen HEUROBOX son distintas, están estrechamente relacionadas y cada una tiene un impacto en las demás. Por ejemplo, la seguridad está estrechamente relacionada con la ergonomía, ya que un robot diseñado ergonómicamente tiene menos probabilidades de causar daño al usuario. Además, la funcionalidad está estrechamente relacionada con las interfaces, ya que las interfaces deben diseñarse para permitir que el robot realice sus tareas previstas de manera eficiente y efectiva. Finalmente, la UX y la aceptación tecnológica se verán afectadas por las cuatro categorías, ya que la seguridad, la ergonomía, la funcionalidad y las interfaces del robot influirán en cómo se percibe la interacción.

- Dependiendo de la complejidad del sistema robótico en sí, el nivel de heurísticas a cumplir puede ser diferente para generar una experiencia satisfactoria. Es decir, un sistema más simple puede cumplir solo con el nivel Básico de heurísticas y generar una experiencia satisfactoria, y un sistema más complejo puede cumplir también con el nivel Básico de heurísticas, pero aún generar una experiencia frustrante para la persona usuaria. Por eso se propone también el Nivel Avanzado, para que sea la persona experta quien elija cuál es el más adecuado para el robot que se va a estudiar. No obstante, se propone realizar una evaluación holística al evaluar la UX en la HRI, de modo que considere las características de la persona usuaria, el contexto y el sistema de manera objetiva y subjetiva, antes, durante y después de la interacción.

En el Capítulo V se presenta el desarrollo y la validación del cuestionario HUROY. A continuación, se resumen las conclusiones principales extraídas de este capítulo:

- Los cuestionarios de UX representan una herramienta valiosa en el campo de la HRI en entornos industriales. A pesar de su potencial, la falta de cuestionarios específicos para evaluar la UX y la aceptación de la tecnología en contextos de HRI industriales es evidente. La creación y validación de un cuestionario como HUROY es un paso esencial hacia la mejora del diseño y la aceptación de sistemas HRI en este contexto.
- La integración de cuestionarios con otras herramientas de evaluación objetiva o cualitativa, tal y como se propone en el modelo ITPX-ROBOTS, ofrece una comprensión más completa de cómo las personas usuarias perciben y aceptan la tecnología en estos entornos.
- El cuestionario HUROY se compone de 7 categorías principales, cada una de las cuales incluye un conjunto de preguntas específicas para un total de 41 preguntas en total. Estas categorías se enfocan en evaluar la percepción de la utilidad del sistema (6 preguntas), la facilidad con la que se puede utilizar (9 preguntas), la sensación de seguridad al usarlo (4 preguntas), la capacidad de controlarlo (8 preguntas), la facilidad de aprender a utilizarlo (3 preguntas), la actitud hacia el sistema (5 preguntas) y la satisfacción general (7 preguntas). Las respuestas a estas preguntas se registran en una escala de 7 puntos de Likert.

En el Capítulo VI se valida el modelo ITPX-ROBOTS mediante dos casos de estudio. A continuación, se resumen las conclusiones extraídas del caso de estudio 1:

- El primer caso de estudio involucra dos procedimientos en los cuales las personas participantes realizaron una tarea de ensamblado a diferentes velocidades y niveles de aceleración. El protocolo de análisis, es decir el modelo ITPX-ROBOTS, ha demostrado ser correcto, y la falta de correlaciones entre las diferentes variables indica que todas las medidas son necesarias.
- El cuestionario UEQ-S no resulta ser el cuestionario más adecuado para comprender mejor los valores perceptuales. El UEQ-S tiene algunas desventajas, como la pérdida de información detallada (Schrepp et al., 2017), ya que no mide la UX por completo como lo hace su versión original UEQ. Sin embargo, también se observa que las preguntas se alejan de la realidad persona-robot. Por lo tanto, existe la necesidad de identificar un cuestionario más adecuado para entornos robóticos industriales. Ante su carencia en la literatura, se procedió a desarrollar y validar el cuestionario HUROY.

- Los resultados indican que velocidades y aceleraciones más altas generalmente conducen a una mayor carga de trabajo y a una menor valencia emocional y *engagement*.
- La duración de la interacción con el robot puede afectar las respuestas emocionales, ya que interacciones más largas tienden a resultar en puntuaciones más bajas de activación e impacto.
- Las personas participantes calificaron más alto la calidad pragmática, la calidad hedónica y la utilidad percibida cuando el robot se movía más rápido. Sin embargo, al mismo tiempo, la alta velocidad también tuvo un efecto negativo en las calificaciones de la fiabilidad y la controlabilidad.
- La velocidad y la aceleración son variables importantes a considerar para optimizar la UX en entornos de trabajo. Los resultados experimentales demuestran que una velocidad demasiado alta influye negativamente en la experiencia al interactuar con el robot, y una velocidad demasiado lenta también la afecta negativamente. Por lo tanto, se ha definido el concepto de “campana de velocidad de HRI” en la Figura 57 que ilustra que una velocidad y aceleración del robot demasiado alta se perciben como inseguras e inconfiables por parte de la persona usuaria, pero una velocidad y aceleración demasiado bajas también generan desconfianza en la persona.
- El orden en que se realizan las tareas influye en las expectativas y respuestas emocionales de las personas participantes hacia el robot. Esto es evidente en las diferencias en las puntuaciones entre los procedimientos LR y RL, donde el orden de las tareas parece impactar en los resultados. Es esencial considerar la influencia de la facilidad de aprendizaje al diseñar la HRI, ya que puede influir en la UX y el rendimiento. Al comprender el efecto del orden de las tareas en las interacciones persona-robot, es posible optimizar el diseño de los robots para satisfacer mejor las necesidades y expectativas de las personas y aumentar su bienestar en la nueva era industrial. El estudio ha revelado una "paradoja de la aprendibilidad" en HRI (Figura 58), donde los tiempos de interacción más largos mejoran el rendimiento y la facilidad de aprendizaje de las personas participantes, pero al mismo tiempo conducen a una disminución en la UX.
- Se han observado diferencias en cómo los hombres y las mujeres perciben y responden a un robot, especialmente en lo que respecta a la valencia emocional, memorización, carga mental y *engagement*. Estas diferencias podrían tener implicaciones importantes para el diseño y uso de robots en diversos entornos. Sin embargo, debido al tamaño de la muestra y la posible influencia de otros factores, se necesita más investigación para comprender y abordar completamente estas diferencias.
- Los resultados del primer caso de estudio proporcionan información valiosa sobre cómo las variables cinemáticas de un robot, como la velocidad y la aceleración, influyen en los factores humanos. Mediante los indicadores de desempeño, se observa que cuanto más rápido se mueve el robot, más pronto se completaba la tarea. Sin embargo, el orden de las tareas influye en los resultados en un estudio de rendimiento de tareas. Esto sugiere que la facilidad de aprendizaje de la interacción mejora el rendimiento.

A continuación, se resumen las conclusiones extraídas del caso de estudio 2:

- El segundo caso de estudio involucra dos procedimientos en los cuales las personas participantes realizaron una tarea de desensamblado con una

interfaz unimodal y con una interfaz multimodal. El modelo ITPX-ROBOTS, ha demostrado ser correcto y aplicable en con diferentes tipos de interfaces.

- Se observa una mejora en los indicadores de desempeño en la tarea que se ejecuta en segundo lugar en ambos procedimientos. Esto sugiere que la experiencia previa con una modalidad de interfaz específica puede mejorar el rendimiento en la segunda ejecución de la tarea, ya que las personas participantes pueden desarrollar estrategias más eficientes y adaptarse mejor.
- El *engagement* es más alto al inicio de la interacción con el robot, pero tiende a disminuir con el tiempo.
- En cuanto a la memorización, se observa que en ambos procedimientos las personas participantes experimentan mayor nivel de memorización al interactuar con la interfaz unimodal. Se puede concluir, por lo tanto, que la interfaz multimodal, independientemente del procedimiento está relacionada con una disminución de los niveles de memorización en comparación con una interfaz unimodal.
- La interfaz multimodal proporciona una valencia emocional más positiva en general.
- La interfaz multimodal tiende a reducir la carga mental percibida por las personas usuarias, lo que podría indicar una experiencia de interacción más eficiente y menos demandante en términos cognitivos en comparación con la interfaz unimodal.
- En lo que respecta a los indicadores perceptuales, se destaca que todos los aspectos evaluados en el cuestionario muestran calificaciones significativamente más favorables cuando las personas participantes interactúan con la interfaz en la segunda posición. Es decir, en el Procedimiento I, la interfaz multimodal recibe una valoración favorable en casi todos los ítems, mientras que en el Procedimiento II, es la interfaz unimodal la que obtiene calificaciones más altas. Esta consistente tendencia sugiere que, al realizar la tarea por segunda vez, las personas participantes experimentan un mayor grado de comodidad y familiaridad con la interfaz, lo que puede haber influido positivamente en sus percepciones y evaluaciones, independientemente de las características específicas de cada interfaz. Cabe destacar que la alta aprendibilidad de la tarea también se refleja en los indicadores de desempeño, lo que respalda aún más la idea de que la familiaridad y la comodidad desempeñan un papel crucial en la mejora de las calificaciones de las personas participantes en la segunda ejecución de la tarea.
- En última instancia, los hallazgos de este estudio indican que, al menos en este entorno particular, las interfaces unimodales y multimodales muestran una adaptación y aceptación similares tanto entre hombres como mujeres, lo que sugiere un prometedor camino hacia la reducción de la brecha digital de género en entornos industriales de HRI.

En ambos casos de estudio, se ha dado especial consideración a la selección de una muestra que garantice la equidad de género, asegurando la representación equitativa de mujeres y hombres. Este enfoque se ha adoptado con el propósito de mitigar posibles sesgos en la investigación, ya que es ampliamente reconocido que en ocasiones las realidades, experiencias y expectativas de un grupo de personas, en este caso, tomando a los hombres como referencia, tienden a erigirse como la norma, lo que podría resultar en hallazgos parciales y no generalizables. En consecuencia, la inclusión de la perspectiva de género en este estudio se ha erigido como un componente fundamental para asegurar una comprensión más completa y precisa de las dinámicas examinadas.

Además, se ha llevado a cabo un análisis descriptivo y estadístico de los datos, desagregándolos por género, con el objetivo de investigar y abordar la brecha digital de género ya existente. En este contexto, se ha observado que, en la tarea de menor velocidad del primer caso de estudio, se manifiesta una diferencia significativa en la tasa de errores, siendo las mujeres quienes han cometido un menor número de fallos. En el segundo caso de estudio no se ha observado ninguna diferencia entre hombres y mujeres. No obstante, es importante destacar que la muestra utilizada en estos casos de estudio consta de 20 personas, distribuidas equitativamente entre 10 hombres y 10 mujeres. Dado este tamaño de muestra limitado, es fundamental reconocer que los hallazgos obtenidos no pueden extrapolarse de manera definitiva a toda la población. Por lo tanto, la necesidad de continuar investigando y analizando la HRI desde una perspectiva de género persiste.

Como conclusión general, esta tesis doctoral ha desarrollado un marco teórico denominado HRIX, que postula que la interacción entre una persona y un robot se ve influida por tres conjuntos de factores: las características de la persona, las características del robot y las características del contexto. Estos elementos en conjunto determinan el grado de ajuste y la calidad de la interacción, lo que se denomina HRI FIT. Este ajuste, a su vez, impacta en diversos componentes de la UX, que se definen en este marco como utilidad, facilidad de uso, comodidad, aprendizaje y control. Estos aspectos de la UX, a su vez, influyen en factores humanos clave, como actitud e intención de uso, confianza, seguridad percibida y satisfacción.

Además, la tesis presenta la taxonomía HUBOXT, que identifica los factores que influyen en la aceptación de la interacción humano-robot por parte del usuario. Este marco teórico y taxonomía sientan las bases para el nuevo modelo propuesto, un modelo holístico que permite evaluar el éxito de la interacción humano-robot en entornos industriales.

En este sentido, se propone un procedimiento que facilita el análisis de la HRI en la realización de tareas y en la percepción de la persona usuaria, teniendo en cuenta medidas objetivas y subjetivas. Este nuevo modelo, denominado ITPX-ROBOTS, proporciona a las empresas manufactureras una herramienta para comprender el impacto de la integración de la HRI. Esto, a su vez, les permitirá optimizar procesos, mejorar la calidad y la productividad. En un contexto en constante transformación de los entornos laborales, donde la adopción de tecnologías es fundamental, este modelo se convierte en una guía esencial para la toma de decisiones en lo que respecta a la incorporación de la HRI en las empresas. Su importancia radica en no perder de vista la apropiación tecnológica y la mejora continua de los lugares de trabajo.

Como se ha mencionado previamente, el modelo se basa en una perspectiva integral que combina aspectos pragmáticos con aquellos relacionados con la experiencia vivida durante la interacción entre la persona usuaria y el robot. Este enfoque abarca la captura de medidas fisiológicas y la percepción de las personas mientras realizan las tareas asignadas. Además, ofrece resultados significativos al incorporar el componente emocional, que desempeña un papel crucial, en conjunto con la evaluación objetiva del desempeño en las tareas. Todo ello mediante herramientas y equipos que permiten recopilar información clave de la interacción de manera imparcial, sin interferir en el proceso de análisis, lo que añade rigor a los resultados obtenidos.

7.2 Validación de las hipótesis

A continuación, se argumenta la validación de cada una de las hipótesis planteadas:

Hipótesis 1

En el contexto de HRI en un entorno industrial, cuanto mayor sea el indicador HEUROBOX incrementará los valores de valencia emocional y engagement, mientras que los valores de memorización y carga mental se verán reducidos.

La validación de esta hipótesis se ha llevado a cabo mediante el cálculo de correlaciones entre el indicador HEUROBOX y los indicadores fisiológicos en los casos de estudio, arrojando resultados que respaldan parcialmente la hipótesis. Los detalles se encuentran en la Tabla 86, donde se muestra el índice de correlación entre el indicador HEUROBOX y los indicadores fisiológicos.

Se ha confirmado la relación negativa (-0,714) entre el indicador HEUROBOX y la carga mental, lo que indica que un mayor cumplimiento de las heurísticas se traduce en una menor carga mental experimentada por las personas usuarias.

Asimismo, se ha verificado que la valencia emocional (0,654) y el *engagement* (0,893) se relacionan positivamente con el indicador HEUROBOX, lo que sugiere que un mayor cumplimiento de los heurísticos se asocia con una mejora en la valencia emocional y el *engagement* de las personas usuarias.

Contrariamente a lo anticipado al formular la hipótesis, se ha observado una correlación positiva (0,855) entre el indicador HEUROBOX y la memorización. Para entender este fenómeno, es crucial considerar el contexto de los casos de estudio y la posible influencia de la complejidad de las tareas en esta relación.

En primer lugar, la investigación se ha basado en dos casos de estudio que involucraban diferentes tareas. Estas tareas no solo variaban en su complejidad, sino que también podrían haber presentado diferencias en la demanda cognitiva que imponían a las personas participantes. Estas variaciones en la naturaleza de las tareas podrían haber influido en los resultados observados.

Una posible explicación de la correlación positiva entre HEUROBOX y la memorización es que HEUROBOX podría no ser lo suficientemente preciso o específico para capturar adecuadamente las características distintivas de cada tarea. Es decir, HEUROBOX podría estar midiendo aspectos generales de la interacción en lugar de considerar las particularidades de las demandas cognitivas específicas de cada tarea. Esto podría haber llevado a una correlación aparente que, en realidad, no refleja la relación deseada entre el indicador HEUROBOX y la memorización.

Tabla 86: Correlación entre el indicador HEUROBOX y los indicadores fisiológicos.

	Memorización	Carga mental	Valencia	Engagement
Indicador HEUROBOX	0,855	-0,714	0,654	0,893

Hipótesis 2

En el contexto HRI en un entorno industrial, cuanto más elevados sean los valores de velocidad y aceleración del robot, incrementará la carga mental, mientras que los valores de la valencia emocional y la percepción general de la experiencia de uso serán disminuidos.

La validación de esta hipótesis se ha realizado mediante el primer caso de estudio, en el cual se ha utilizado el modelo ITPX-ROBOTS para evaluar el impacto de las variables cinemáticas en los factores humanos.

Como se refleja en la Tabla 65 y se deduce de los resultados obtenidos en este caso de estudio, se aprecia que el aumento de las variables cinemáticas conlleva un incremento en la carga mental y una disminución en la valencia emocional. Es importante destacar que, en contraposición a la hipótesis, cuando las variables cinemáticas son demasiado bajas, la percepción de la experiencia de uso también se ve afectada negativamente, lo que no se había previsto inicialmente. Se encontró que las calificaciones en la calidad pragmática, la calidad hedónica y la utilidad percibida eran más altas cuando las variables cinemáticas eran mayores.

Esta observación ha llevado a la formulación de lo que se denomina la "campana de velocidad en HRI" en la Figura 57. Esta idea sugiere que no solo las dinámicas excesivamente altas, sino también las demasiado bajas, pueden generar un efecto negativo en la percepción y la experiencia de las personas usuarias. Por lo tanto, se enfatiza la importancia de encontrar un equilibrio óptimo en las variables cinemáticas utilizadas en la HRI en entornos industriales.

En resumen, la Hipótesis 2 se valida parcialmente, ya que se confirma que cuando mayores son las variables cinemáticas, incrementa la carga mental y disminuye la valencia emocional. Sin embargo, se revela un impacto negativo cuando las variables cinemáticas son demasiado bajas en la percepción de la experiencia de uso.

Hipótesis 3

En el contexto de HRI en un entorno industrial, el uso de interfaces multimodales conlleva una disminución en los valores de memorización y carga mental y un incremento de la experiencia de uso percibida en comparación con el uso de una interfaz unimodal.

Mediante el análisis del segundo caso de estudio, se confirma la validez de la hipótesis 3. Se evidencia que tanto en el Procedimiento I como en el Procedimiento II, es decir, independientemente del orden de uso, los niveles de memorización y carga mental resultan inferiores al emplear la interfaz multimodal en contraste con la interfaz unimodal. Es relevante destacar que esta validación se aprecia tanto en participantes de género masculino como femenino, lo que confiere solidez a los resultados. Además, es notable que la mejora no se limita únicamente a aspectos cognitivos y de carga mental, sino que también se refleja en una percepción general más positiva de la experiencia de uso, según lo indicado en los resultados del cuestionario HUOX.

7.3 Aportaciones más relevantes

Dentro del contexto de esta tesis doctoral, se destacan las siguientes contribuciones más relevantes:

Publicaciones en revistas:

1. Apraiz, A., Lasa, G. & Mazmela, M. Evaluation of User Experience in Human-Robot Interaction: A Systematic Literature Review. *International Journal of Social Robotics* (2023). <https://doi.org/10.1007/s12369-022-00957-z> (I.F.:1,18, Q1)
2. Apraiz, A., Mulet Alberola, J.A., Lasa, G., Mazmela, M., & Ngoc Nguyen, H. Development of a new set of Heuristics for the evaluation of Human-Robot Interaction in industrial settings: Heuristics Robots Experience (HEUROBOX). *Frontiers in Robotics and AI* (2023). AI 10:1227082. <https://doi.org/10.3389/frobt.2023.1227082> (I.F.: 0,73, Q2)
3. Apraiz, A., Lasa, G., Montagna, F., Blandino, G., Triviño-Tonato, E. & Dacal-Nieto, A. An Experimental Protocol for Human Stress Investigation in Manufacturing Contexts: Its Application in the NO-STRESS Project. *Systems* (2023), 11, 448. <https://doi.org/10.3390/systems11090448> (I.F.: 0,48, Q2)
4. Apraiz-Iriarte, A., Lasa-Erle, G., Mazmela-Etxabe, M. Application of questionnaires for evaluation of user experience in industrial environments. *DYNA* (2021), 96(3). 235. <https://doi.org/10.6036/10025> (I.F.: 0,16, Q4)
5. Apraiz-Iriarte, A., Lasa-Erle, G., Mazmela-Etxabe, M. Present and future of physiological monitoring to evaluate user experience. *DYNA* (2021), 96(4). 342. <https://doi.org/10.6036/10218> (I.F.: 0,16, Q4)
6. Apraiz-Iriarte, A., Lasa-Erle, G., Mazmela-Etxabe, M. Evaluating user experience with physiological monitoring: a systematic literature review. *DYNA New Technologies* (2021), 8(1). [20 p.]. <https://doi.org/10.6036/NT10072>
7. Apraiz-Iriarte, A., Lasa-Erle, G. Questionnaires for evaluation of user experience and usability in industrial environments: bibliographic review. *DYNA New Technologies* (2020), 7(1). [21 p.]. <https://doi.org/10.6036/NT9845>

Publicaciones en congresos:

1. Apraiz, A., Lasa, G., Mazmela, M., Escallada, O., & González de Heredia, A. (2023). Development and validation of Human-Robot eXperience (HUROX) questionnaire for industrial collaborative contexts. 27th International Congress on Project Management and Engineering. Donostia (Spain).
2. Apraiz Iriarte, A., Lasa Erle, G., Serrano Muñoz, A., Elguea Aguinaco, Í., & Arana Arexolaleiba, N. (2022). Evaluation of the user experience of an industrial robotic environment in virtual reality. 26th International Congress on Project Management and Engineering. Terrassa (Spain).
3. Apraiz Iriarte, A., Lasa, G., & Mazmela Etxabe, M. (2022). Evaluación heurística de la interacción persona-robot en entornos industriales. Congreso Internacional de Interacción Persona-Ordenador (Interacción). Teruel (Spain).
4. Apraiz Iriarte, A., Aztiria Mendizabal, A., & Lasa Erle, G. (2021). Robot-aren itxura estetikoak eta erabiltzaileen preferentziak. *Ikergazte*. Vitoria-Gasteiz (Spain). <https://dx.doi.org/10.26876/ikergazte.iv.03.05>
5. Apraiz Iriarte, A., Lasa Erle, G., & Mazmela Etxabe, M. (2021) User preferences and associations with light or dark interfaces. 25th International Congress on Project Management and Engineering. Alcoy (Spain).

Por último, es relevante mencionar que este trabajo ha contribuido al desarrollo y publicación de los siguientes trabajos, en los cuales se participa como coautora:

1. Mazmela-Etxabe, M., Lasas-Erle, G., Apraiz-Iriarte, A. Interface design and its impact on software appropriation. DYNA (2021), 96(1), 16. <http://dx.doi.org/10.6036/9939> (I.F.: 0,16, Q4)
2. Mazmela-Etxabe, M., Lasas-Erle, G., Apraiz-Iriarte, A. Multimethod approach to measure the design factor impact on the appropriation of a competitive intelligence interface. DYNA New Technologies (2020), 7(1). [13 p.]. <https://doi.org/10.6036/NT9805>

7.4 Limitaciones de la investigación

Es fundamental considerar ciertas limitaciones que podrían influir en la interpretación y generalización de los hallazgos. Estas limitaciones aportan un contexto en el cual los resultados deben ser evaluados y aplicados de manera adecuada.

La principal limitación que debe destacarse en esta investigación se refiere al número limitado de casos de estudio utilizados para validar el modelo ITPX-ROBOTS. Aunque los dos casos de estudio realizados se consideran completos y han contribuido significativamente a la comprensión del modelo, es importante reconocer que la inclusión de un mayor número de casos de estudio habría aportado beneficios adicionales.

A pesar de la sólida base teórica, la realización de únicamente dos casos de estudio puede generar ciertas limitaciones en términos de generalización de los resultados. Dado que los casos de estudio pueden variar en función de múltiples factores, como el entorno, las condiciones específicas y las interacciones particulares entre personas y robots, contar con una muestra más amplia de casos habría permitido una validación más sólida y una comprensión más completa de la aplicabilidad del modelo en diferentes contextos.

Asimismo, un mayor número de casos de estudio podría haber proporcionado una mayor diversidad de perspectivas y escenarios, lo que habría enriquecido la evaluación del modelo y permitido la identificación de posibles patrones o tendencias emergentes en la interacción persona-robot en entornos industriales.

A continuación, se presentan las demás limitaciones identificadas con el propósito de enriquecer la comprensión y el valor de las conclusiones alcanzadas. A continuación, se listan las limitaciones habidas en el estado del arte:

- La metodología SLR no garantiza que se identifiquen todas las publicaciones relacionadas con un ámbito de investigación determinado (Aleu & van Aken, 2016). Al haber limitado la búsqueda a artículos revisados por pares, es posible que se hayan omitido estudios de casos publicados en congresos que podrían haber sido relevantes para el estudio. Consciente de estas limitaciones, se ha realizado un esfuerzo continuo para ofrecer una visión completa y actualizada de la literatura en este campo de investigación. Es relevante señalar que, a pesar de la SLR inicial, se incorporó un subcapítulo adicional que incluye artículos más recientes, con el fin de mantener la información actualizada y abordar la limitación de la cobertura temporal de la SLR.
- El sesgo del revisor: A pesar de haber intentado objetivar la revisión, se podría haber introducido sesgos en algunos casos.
- La elección de las bases de datos utilizadas. Aunque seleccionamos estratégicamente las bases de datos para asegurar una cobertura adecuada de esta área de investigación y designamos una estrategia de búsqueda para

asegurar que se captaba el mayor número posible de publicaciones, es posible que, si hubiéramos utilizado alguna base de datos adicional, podríamos haber identificado más artículos significativos para la re-búsqueda.

- Los criterios de EC. Si se hubiesen definido otras preguntas de EC, el resultado de la SLR habría sido completamente distinto. Sin embargo, el objetivo era centrarse en los factores humanos, el género y la accesibilidad. Lo mismo ocurre con el hecho de haber fijado el límite de la EC en 7. Si se hubiese fijado un valor más bajo, tal vez se habrían identificado otros artículos relevantes. No obstante, tanto las preguntas de la EC utilizadas como el valor de corte han proporcionado artículos de calidad, lo que se ve respaldado por la caracterización de la literatura. Ya que el 96% de los artículos corresponden a artículos publicados en revistas indexadas, el 29% en el primer cuartil y el 56% en el segundo cuartil.
- Otra limitación es la restricción de los resultados únicamente a los idiomas inglés y español.

A continuación, se listan las limitaciones habidas en el desarrollo y validación de HEUROBOX:

- La selección de las bases de datos. La elección de las bases de datos fue un proceso estratégico para asegurar una cobertura integral del área de investigación. A pesar de haber diseñado una estrategia de búsqueda para capturar publicaciones relevantes, es posible que la exclusión de bases de datos adicionales haya resultado en la omisión de artículos significativos que podrían haber contribuido a los resultados de la investigación.
- Sesgo en la selección de heurísticas. La selección y agrupación de las heurísticas se basaron en la interpretación de la literatura por parte de los investigadores, lo que podría haber introducido un sesgo en la selección y agrupación de las heurísticas, a pesar de los esfuerzos realizados para evitarlo.

A continuación, se enumeran las limitaciones identificadas en el desarrollo y validación del cuestionario HUOX:

- Selección de los cuestionarios ejemplares: La elección de los cuestionarios ejemplares fue un proceso estratégico destinado a garantizar una cobertura integral del área de investigación. A pesar de haber seleccionado cuidadosamente los cuestionarios más relevantes, es posible que la exclusión de otros cuestionarios haya resultado en la omisión de ítems significativos que podrían haber influido en los resultados de la investigación.
- Sesgo en la selección de los ítems: La selección y agrupación de los ítems se basaron en la interpretación de la literatura por parte de los investigadores, lo que podría haber introducido un sesgo en la elección y agrupación de las heurísticas, a pesar de los esfuerzos realizados para evitarlo.
- Sesgo en la evaluación por parte de expertos: A pesar de haber seleccionado a personas expertas en diversas temáticas influyentes y relevantes para la elaboración del cuestionario, existe la posibilidad de que estas personas hayan introducido un sesgo al evaluar cada uno de los ítems que finalmente conformaron el cuestionario.

A continuación, se listan las limitaciones de los casos de estudio:

- Tamaño de la Muestra: En los estudios de J. R. Lewis (1994), Nielsen (2000) y Virzi (1992) afirman que 5 participantes son suficientes para identificar el 80% de los problemas de usabilidad. No obstante, tamaños de muestra reducidos pueden resultar en una alta variabilidad en los resultados de las pruebas que

no puede ser completamente ajustada (Cazañas et al., 2017). Como concluimos en estudios previos (Apraiz et al., 2021), no existe un tamaño de muestra ampliamente validado para pruebas de UX con monitoreo fisiológico. Por lo tanto, el tamaño de muestra en este experimento podría ser demasiado pequeño para obtener las posibles correlaciones. Ambos estudios han sido conducidos con una muestra relativamente reducida de participantes (20 personas en cada caso de estudio, lo que, al considerar dos procedimientos, resulta en 10 participantes por procedimiento). Esto podría limitar la generalización de los resultados a poblaciones más grandes. Sin embargo, a pesar de que el tamaño de la muestra sea limitado, los estudios proporcionan una visión inicial valiosa sobre el impacto de las variables cinemáticas y las interfaces unimodales y multimodales en un entorno industrial. Los hallazgos pueden servir como base para investigaciones futuras con muestras más amplias.

- Contexto específico de estudio: Es esencial señalar que el escenario de investigación se ha circunscrito a un ámbito industrial particular, involucrando una tarea de ensamblado en el primer caso de estudio y de desensamblado en el segundo. Además, ambos estudios compartieron el mismo robot como sujeto de análisis, si bien presentando distinto comportamiento debido a su programación diferenciada (en el primer caso de estudio la programación del robot era convencional y en el segundo basada en aprendizaje por refuerzo). A pesar de que el contexto sea específico, los resultados aún ofrecen información relevante sobre cómo las variables cinemáticas y las interfaces unimodales y multimodales pueden influir en la HRI en entornos colaborativos. Los estudios, aun dentro de su contexto acotado, se traducen en un avance hacia la comprensión de la dinámica interactiva entre humanos y robots en un entorno industrial.
- Sesgo en la configuración de los casos de estudio. Las tareas en los casos de estudio se simplificaron para investigar cómo las variables cinemáticas y las interfaces unimodales y multimodales influyen en la interacción. En consecuencia, ningún participante podría haber enfrentado dificultades notables al llevar a cabo el ensamblado o desensamblado. No obstante, esta simplificación podría haber introducido un sesgo en los casos de estudio, ya que las tareas podrían haber sido percibidas como excesivamente sencillas, lo que podría haber disminuido su realismo para las personas participantes.
- Aplicabilidad de la configuración propuesta. Los dispositivos propuestos en esta investigación para la evaluación de factores fisiológicos demuestran su aplicabilidad en un entorno de laboratorio, sin embargo, podrían presentar limitaciones en un contexto real. Dichos dispositivos restringen en cierta medida el movimiento del participante y no permiten que la persona usuaria se comporte de manera natural como lo haría en una situación real. Además, es necesario considerar el grado de intrusividad que estos dispositivos pueden conllevar en un entorno más práctico.

7.5 Líneas futuras de la investigación

En el marco de esta tesis doctoral, se han alcanzado satisfactoriamente los objetivos previamente establecidos. No obstante, los resultados de la investigación han abierto la puerta a la identificación de nuevas áreas de estudio que merecen una exploración detallada en un futuro. En consecuencia, en esta sección se presentan las siguientes direcciones a seguir en el futuro:

7.5.1 Optimizar la herramienta HEUROBOX

En el contexto de la investigación futura, se plantea la optimización de la herramienta HEUROBOX. Aunque HEUROBOX es una herramienta basada en heurísticas que se destaca por su facilidad de aplicación y bajo costo, todavía existen oportunidades para mejorar su capacidad de evaluación de la HRI.

7.5.1.1 *Contemplar la multimodalidad de manera más efectiva*

Como parte de las investigaciones futuras, se plantea la necesidad de abordar de manera más efectiva la multimodalidad en la herramienta HEUROBOX. HEUROBOX ya ofrece apartados concretos para evaluar cada una de las interfaces, incluyendo opciones visuales, basadas en voz, gestuales, y de lenguaje natural, así como interacciones físicas y hápticas.

En relación a las interacciones multimodales, que involucran la combinación de múltiples modalidades sensoriales para lograr una comunicación más natural con los robots, existe un potencial considerable para crear interfaces de usuario y canales de interacción altamente efectivos. Sin embargo, es importante destacar que la lista de heurísticas propuesta puede generar cierta superposición en las subcategorías relacionadas con las interfaces en la sección de HEUROBOX. Esta superposición podría llevar a la percepción de que un robot con dos tipos de interfaces cumple un menor porcentaje de heurísticas en comparación con un robot que solo tiene una interfaz.

No obstante, es esencial reconocer que cuando estas dos interfaces se complementan de manera eficaz, la calidad general de la interacción puede mejorar significativamente. Por lo tanto, en futuras investigaciones, se requerirá un enfoque más detallado para evaluar las interfaces en un contexto multimodal. Esto implicaría analizar minuciosamente la superposición entre las heurísticas, considerando cómo las diferentes modalidades interactúan entre sí y cómo contribuyen a la UX en su conjunto.

7.5.1.2 *Validar y correlacionar el nivel de cumplimiento con la aceptación del sistema*

Como parte de la investigación futura, se plantea la validación y correlación del indicador HEUROBOX con el nivel de satisfacción y aceptación percibido por las personas usuarias durante su interacción con robots en entornos industriales. Además, se considera relevante explorar la posibilidad de establecer un sistema de clasificación o estándar basado en la complejidad, características, funcionalidades y objetivos específicos de cada tipo de robot.

Este sistema de clasificación permitiría determinar el nivel de cumplimiento de las heurísticas que debe alcanzar cada categoría de robot, proporcionando así a los profesionales que utilizan HEUROBOX una guía coherente y completa para el diseño y la evaluación de robots. Esto, a su vez, fomentaría una interacción más efectiva y eficiente entre las personas y los robots en entornos industriales.

Es importante destacar que la creación de esta clasificación o norma debe ir de la mano de la correlación del cumplimiento de las heurísticas con la satisfacción y aceptación percibidas por parte de las personas usuarias. Esta correlación puede variar según el tipo de robot en cuestión. Por ejemplo, un robot con objetivos más simples podría cumplir un porcentaje menor de las actuales heurísticas y aún ser aceptado por las personas usuarias. En contraste, un robot con objetivos y funcionalidades más complejas podría tener un indicador HEUROBOX más alto, pero no ser aceptado adecuadamente. Por lo tanto, la clasificación propuesta debe estar respaldada por una correlación efectiva para adaptarse a las necesidades y expectativas de las personas usuarias en función de la naturaleza específica de cada robot.

7.5.1.3 La escalabilidad de HEUROBOX

Una línea de investigación futura es la escalabilidad de HEUROBOX. A medida que la HRI continúa avanzando y se identifican nuevos factores y desafíos en esta área, se presenta una oportunidad valiosa para mejorar y expandir iterativamente HEUROBOX para mantenerse al día con los desarrollos emergentes.

La organización estructurada y la clasificación por secciones de HEUROBOX permiten su mejora continua de manera sencilla. Por ejemplo, se podría considerar la inclusión de una sección adicional dedicada a consideraciones éticas en la evaluación de la HRI. Dado que la ética desempeña un papel cada vez más relevante en la adopción y aceptación de la tecnología robótica en diversos contextos, su incorporación en HEUROBOX permitiría abordar de manera más integral los aspectos éticos relacionados con la interacción entre personas y robots en entornos industriales.

Además, a medida que surjan nuevos desafíos y tendencias en la HRI, HEUROBOX podría adaptarse para incluir categorías adicionales o heurísticas específicas que aborden estos temas emergentes. Esto aseguraría que la herramienta siga siendo relevante y efectiva en la evaluación de la robótica industrial. La escalabilidad de HEUROBOX no solo permitiría una evaluación más completa y actualizada de la HRI, sino que también contribuiría a su continua mejora y aplicabilidad en una amplia gama de entornos y contextos industriales en constante evolución.

7.5.2 Validar la fiabilidad de HUOX

Para asegurar la alta fiabilidad del cuestionario, sería pertinente su aplicación en diversos contextos, seguida del cálculo de métricas de fiabilidad. Este proceso de validación comprende la ejecución de pruebas piloto y la realización de análisis estadísticos diseñados para evaluar la coherencia interna y la consistencia de las respuestas. La fiabilidad del cuestionario es de una importancia crítica para garantizar la precisión y consistencia de los datos recopilados.

7.5.3 Identificación de las variables cinemáticas óptimas para la HRI en entornos industriales y la robótica afectiva

Las conclusiones extraídas del primer caso de estudio han arrojado luz sobre la importancia de las variables cinemáticas en la HRI industrial. La formulación de la "campana de velocidad en HRI" en la Figura 57 resalta la necesidad de encontrar un equilibrio óptimo en estas variables, ya que tanto valores demasiado altos como demasiado bajos pueden afectar negativamente la percepción y la experiencia de las personas usuarias.

En este sentido, resulta necesario explorar más a fondo cómo se pueden definir y ajustar estas variables cinemáticas para optimizar la HRI en entornos industriales. Además, es importante considerar cómo estas variables pueden adaptarse a diferentes tareas y situaciones específicas, ya que la complejidad de las tareas y las demandas cognitivas pueden variar significativamente.

En este sentido, las investigaciones futuras podrían centrarse en desarrollar modelos y algoritmos que permitan la regulación dinámica de las variables en tiempo real, en función de las necesidades y preferencias de las personas usuarias. Asimismo, se podría explorar cómo estas variables influyen en otros aspectos de la interacción, como la eficiencia de la tarea, la seguridad y la satisfacción de la persona usuaria. Esta perspectiva abre la puerta a la robótica afectiva, donde el comportamiento del robot se adapta de manera inteligente y empática a las emociones y necesidades de la persona usuaria, lo que podría revolucionar aún más la HRI en entornos industriales.

7.5.4 Medición de la UX Acumulativa y estudios longitudinales

La evaluación de la UX acumulativa se presenta como un enfoque de interés en la investigación actual, con el propósito de comprender la evolución y el cambio en la percepción y satisfacción de las personas usuarias a lo largo del tiempo. La Figura 75 ejemplifica cómo podría adaptarse un diseño experimental para abordar distintos intervalos temporales relacionados con la UX.

Para obtener una comprensión exhaustiva de la interacción inicial del usuario con el sistema, es conveniente emplear métodos como cuestionarios o entrevistas durante el primer encuentro. Estos métodos permiten capturar de manera precisa las impresiones iniciales y las percepciones que emergen al interactuar por primera vez con el producto o sistema.

La UX momentánea, ya se evalúa de forma precisa empleando el modelo ITPX-ROBOTS.

La medición de la UX acumulativa entraña la repetición del estudio experimental en múltiples ocasiones a lo largo del tiempo, involucrando al mismo grupo de participantes en diferentes sesiones. Este enfoque posibilita la comparación de los resultados obtenidos en diversas instancias, lo que a su vez permite identificar patrones y cambios en la experiencia del usuario a medida que se familiarizan con el sistema y acumulan experiencia en su uso. Es importante considerar que, para medir la UX acumulativa de manera efectiva, se deben tener en cuenta factores como el número de sesiones, el intervalo entre cada sesión y la consistencia en la metodología utilizada en cada repetición del estudio.

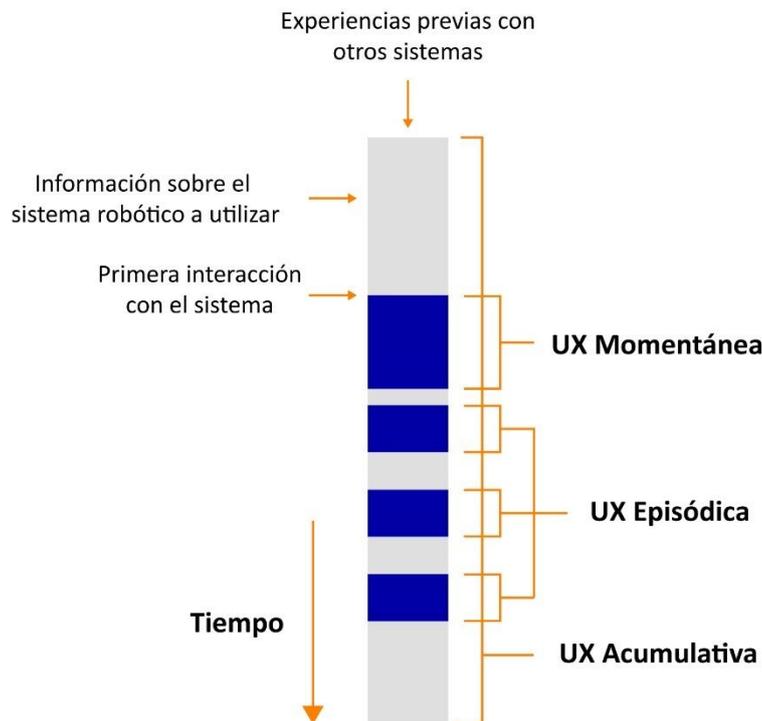


Figura 75: Aplicación posible en un estudio longitudinal.

Por lo tanto, se presenta una dirección futura de investigación que podría enfocarse en la realización de un estudio longitudinal centrado en la aceptación y la UX en entornos industriales con robots. Esta perspectiva no solo permitiría evaluar la respuesta inicial de los trabajadores ante la incorporación de robots colaborativos en su entorno laboral, sino también comprender cómo estas percepciones evolucionan y se modifican a lo largo del tiempo. En este estudio longitudinal, se podrían llevar a cabo mediciones periódicas de la aceptación de la

HRI y la calidad de UX a lo largo de diferentes etapas del proceso de implementación y adopción de la tecnología mediante el modelo ITPX-ROBOTS. Esto permitiría analizar si ciertas variables o factores humanos que influyen en la aceptación y la UX cambian. Por ejemplo, se podría investigar si la seguridad percibida aumenta a medida que las personas trabajadoras ganan experiencia y confianza en la tecnología, o si la importancia de la adaptabilidad del robot a las demandas físicas y cognitivas cambia a medida que las personas trabajadoras se familiarizan con las capacidades del sistema.

Además, este enfoque longitudinal permitiría identificar patrones de resistencia o aceptación sostenida en diferentes grupos de trabajadores, teniendo en cuenta variables como la experiencia tecnológica, la posición laboral o características personales. A medida que se recopilen datos a lo largo del tiempo, podrían desarrollarse estrategias más efectivas para gestionar el cambio organizacional, abordar preocupaciones específicas y maximizar la aceptación y satisfacción de los trabajadores en entornos de HRI industriales en constante evolución.

7.5.5 Análisis de factores contextuales en casos de estudio más diversificados

Una línea futura de investigación sugerida es la exploración de cómo los resultados obtenidos en los casos de estudio actuales pueden aplicarse a una gama más amplia de contextos industriales y tareas específicas. Si bien los casos de estudio actuales proporcionaron información valiosa sobre la HRI en un entorno de ensamblado y desensamblado, es importante considerar que diferentes industrias y tipos de tareas pueden presentar desafíos y dinámicas únicas en la interacción.

Se podría llevar a cabo una investigación más amplia que abarque diversos sectores industriales, como la logística, la electrónica o la industria alimentaria. Cada uno de estos sectores podría involucrar tareas específicas y requerimientos de interacción distintos. Esto nos lleva a que el modelo completo y concretamente la herramienta de evaluación heurística HEUROBOX, necesitarían algunas modificaciones de cara a poder adaptarlo a contextos de uso más amplios.

Otro aspecto importante a considerar en futuras investigaciones es la estructura del equipo persona-robot. ¿Cómo interactúan las personas operarias humanos con múltiples robots en un entorno de trabajo colaborativo? ¿Cómo influye esto en la eficiencia y en la percepción de la tarea? Estas preguntas podrían abordarse a través de casos de estudio que simulen situaciones más realistas de interacción entre humanos y múltiples robots en diversos roles y responsabilidades.

7.5.6 La ética y la accesibilidad en la HRI

Un campo emergente y de creciente relevancia en la HRI es la ética. A medida que la colaboración entre personas y robots se vuelve más común en diversos entornos, es esencial abordar las cuestiones éticas que surgen en esta dinámica. Si bien los robots pueden aportar muchos beneficios, también pueden plantear desafíos éticos.

En este sentido, una línea de investigación podría abordar preguntas fundamentales, como la privacidad y seguridad de los datos y personas en la HRI. Además, se podría investigar la toma de decisiones éticas en situaciones donde los robots y las interfaces pueden influir en la seguridad y el bienestar de las personas trabajadoras. La toma de decisiones éticas en situaciones en las cuales los robots y las interfaces tecnológicas tienen la capacidad de influir en la seguridad y el bienestar de las personas trabajadoras. La introducción de robots en el entorno laboral puede plantear dilemas éticos, como la asignación de responsabilidades en caso de accidentes o la toma de decisiones críticas durante situaciones de emergencia. Investigar cómo desarrollar marcos éticos y protocolos de toma de decisiones adecuados en la

HRI industrial es esencial para garantizar que las interacciones entre personas y robots sean seguras, justas y éticas.

Otro aspecto importante reside en el análisis de la equidad y la justicia en la colaboración entre personas y robots en el entorno laboral. Se plantean interrogantes esenciales, como la forma en que se pueden diseñar interfaces para prevenir la aparición de desigualdades en el lugar de trabajo. Además, se busca la manera de asegurar que la implementación de tecnologías de interfaz no conduzca a la marginación de ciertos trabajadores o grupos. Este análisis debería buscar no solo abordar la equidad y la justicia en la colaboración persona-robot, sino también garantizar que la introducción de tecnologías mediante las interfaces sea inclusiva y promueva un ambiente de trabajo equitativo y accesible para todas las personas empleadas, sin importar sus características individuales.

Dentro de esta línea de investigación, se plantea la posibilidad de investigar en profundidad la percepción y aceptación de las interfaces desde una perspectiva ética. Se busca entender cómo la percepción de las personas trabajadoras sobre la ética de la interacción puede influir en su compromiso con las tareas en entornos de HRI. Esto implica analizar cómo las consideraciones éticas y morales que rodean la interacción con robots pueden afectar la actitud y el rendimiento de las personas usuarias en el contexto laboral.

En última instancia, una investigación en ética y HRI podría contribuir en gran medida a establecer pautas y mejores prácticas para la implementación de interfaces en entornos industriales. Esto no solo garantizaría una colaboración eficiente y productiva, sino también una colaboración ética y equitativa que priorice el bienestar de las personas trabajadoras y promueva valores fundamentales en la HRI.

Bibliografía

8 BIBLIOGRAFÍA

- Adamides, G. (2020). Heuristic Evaluation of the User Interface for a Semi-Autonomous Agricultural Robot Sprayer. *AGRIS on-line Papers in Economics and Informatics*, 12(665-2022-407), 3-12.
- Adamides, G., Christou, G., Katsanos, C., Xenos, M., & Hadzilacos, T. (2014). Usability guidelines for the design of robot teleoperation: A taxonomy. *IEEE Transactions on Human-Machine Systems*, 45(2), 256-262.
- Akalin, N., Kristoffersson, A., & Loutfi, A. (2019). Evaluating the sense of safety and security in human-robot interaction with older people. *Social robots: Technological, societal and ethical aspects of human-robot interaction*, 237-264.
- Akalin, N., Kristoffersson, A., & Loutfi, A. (2022). Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures. *International journal of human-computer studies*, 158, 102744. <https://doi.org/10.1016/j.ijhcs.2021.102744>
- Aladwani, A. M., & Palvia, P. C. (2002). Developing and validating an instrument for measuring user-perceived web quality. *Information & management*, 39(6), 467-476.
- Aleu, F. G., & van Aken, E. M. (2016). Systematic literature review of critical success factors for continuous improvement projects. *International Journal of Lean Six Sigma*.
- Aliagas Ocaña, I. (2022). *Análisis neuropsicofisiológico de la eficacia del emplazamiento de producto en videojuegos*. Universidad Complutense de Madrid.
- Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality progress*, 40(7), 64-65.
- Almeida, L., Menezes, P., & Dias, J. (2020). Interface Transparency Issues in Teleoperation. *Applied Sciences*, 10(18), 6232. <https://doi.org/10.3390/APP10186232>
- Alzoubi, O., D'Mello, S. K., & Calvo, R. A. (2012). Detecting naturalistic expressions of nonbasic affect using physiological signals. *IEEE Transactions on Affective Computing*, 3(3), 298-310. <https://doi.org/https://doi.org/10.1109/T-AFFC.2012.4>
- Andonovski, B., Ponsa, P., & Casals, A. (2010). Towards the development of a haptics guideline in human-robot systems. *3rd International Conference on Human System Interaction*, 380-387.
- Apraiz, A., & Lasa, G. (2020). Questionnaires for the evaluation of user experience and usability in industrial environments: A bibliographic review. *DYNA New Technologies*, 7(1), [21 p.]-[21 p.]. <https://doi.org/10.6036/nt9845>
- Apraiz, A., Lasa, G., & Mazmela, M. (2021). Evaluating User Experience with Physiological monitoring: A Systematic Literature Review. *DYNA New Technologies*, 8(1), 21-undefined. <https://doi.org/https://doi.org/10.6036/NT10072>
- Apraiz, A., Lasa, G., & Mazmela, M. (2023). Evaluation of User Experience in Human-Robot Interaction: A Systematic Literature Review. *International Journal of Social Robotics*, 1-24. <https://doi.org/10.1007/S12369-022-00957-Z/TABLES/13>
- Apraiz, A., Lasa, G., Montagna, F., Blandino, G., Triviño-Tonato, E., & Dacal-Nieto, A. (2023). An Experimental Protocol for Human Stress Investigation in Manufacturing Contexts: Its Application in the NO-STRESS Project. *Systems*, 11(9), 448. <https://doi.org/10.3390/systems11090448>

- Apraiz, A., Lasa, G., Serrano, A., Elguea, Í., & Arana-Arexolaleiba, N. (2022). Evaluation of the User Experience of an Industrial Robotic Environment in Virtual Reality. *International Congress on Project Management and Engineering*, 7.
- Apraiz, A., Mulet Alberola, J. A., Lasa, G., Mazmela, M., & Nguyen Ngoc, H. (2023). *Experts' Evaluation: Datasets of Rating Questionnaires for the HEUROBOX Heuristics to Evaluate Industrial Collaborative Robotic Environments* (Vol. 1). Mendeley Data. <https://doi.org/10.17632/T2NWPWWSMG.1>
- Arai, T., Kato, R., & Fujita, M. (2010). Assessment of operator stress induced by robot collaboration in assembly. *CIRP Annals*, 59(1), 5-8. <https://doi.org/10.1016/J.CIRP.2010.03.043>
- Aranburu, E., Lasa, G., Gerrikagoitia, J. K., & Kepa, J. (2019). *HEMEI: new user experience evaluation tool for Human Machine Interfaces*. <https://doi.org/10.1145/3335595.3335621>
- Aranburu, E., Lasa, G., Gerrikagoitia, J. K., & Mazmela, M. (2020). Case Study of the Experience Capturer Evaluation Tool in the Design Process of an Industrial HMI. *Sustainability*, 12(15), 6228. <https://doi.org/10.3390/su12156228>
- Aranburu Zabalo, E. (2020). *Demox: Nuevo modelo de trabajo de diseño y evaluación de la UX en interfaces industriales* [Tesis doctoral, Mondragon Unibertsitatea]. <http://hdl.handle.net/20.500.11984/1876>
- Aranburu Zabalo, E., Lasa Erle, G., Reguera Bakhache, D., Gerrikagoitia, J. K., & Iruretagoiena Irazusta, G. (2017). Metodología UCAD: Nuevo procedimiento de diseño de interfaces centrado en el usuario para la industria 4.0. *Dyna new technologies*, 4(1), 15-p.
- Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., & Sharples, S. (2021). Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. *International Journal of Human Computer Studies*, 145, 102522. <https://doi.org/10.1016/j.ijhcs.2020.102522>
- Aromaa, S., Frangakis, N., Tedone, D., Viitaniemi, J., & Aaltonen, I. (2018). Digital Human Models in Human Factors and Ergonomics Evaluation of Gesture Interfaces. *Proceedings of the ACM on Human-Computer Interaction*, 2(EICS). <https://doi.org/10.1145/3229088>
- Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79-86. <https://doi.org/10.1177/0748175613513808>
- Bai, A., & Fuglerud, K. S. (2018). Method for Semi-Automated Evaluation of User Experience Using Brain Activity. *Studies in health technology and informatics*, 256, 811-820. <https://doi.org/https://doi.org/10.3233/978-1-61499-923-2-811>
- Baig, M. Z., & Kavakli, M. (2019). A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. *Multimodal Technologies and Interaction*, 3(2), 37. <https://doi.org/https://doi.org/10.3390/mti3020037>
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4(3), 114-123.
- Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges? A critical analysis of empirical studies of User Experience. *Conference on Human Factors in Computing Systems - Proceedings*, 2689-2698. <https://doi.org/https://doi.org/10.1145/1978942.1979336>

- Bargas-Avila, J. A., Lötscher, J., Orsini, S., & Opwis, K. (2009). Intranet satisfaction questionnaire: Development and validation of a questionnaire to measure user satisfaction with the Intranet. *Computers in Human Behavior*, 25(6), 1241-1250. <https://doi.org/10.1016/J.CHB.2009.05.014>
- Barral, O., Kosunen, I., & Jacucci, G. (2017). No need to laugh out loud: Predicting humor appraisal of comic strips based on physiological signals in a realistic environment. *ACM Transactions on Computer-Human Interaction*, 24(6), 1-29. <https://doi.org/https://doi.org/10.1145/3157730>
- Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2008). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics 2008 1:1*, 1(1), 71-81. <https://doi.org/10.1007/S12369-008-0001-3>
- Baskaran, S., Niaki, F. A., Tomaszewski, M., Gill, J. S., Chen, Y., Jia, Y., Mears, L., & Krovi, V. (2019). Digital Human and Robot Simulation in Automotive Assembly using Siemens Process Simulate: A Feasibility Study. *Procedia Manufacturing*, 34, 986-994. <https://doi.org/10.1016/J.PROMFG.2019.06.097>
- Bauer, A., Gonsior, B., Wollherr, D., & Buss, M. (2009). Heuristic rules for human-robot interaction based on principles from linguistics-asking for directions. *AISB Convention-Symposium on New Frontiers in Human-Robot Interaction*, 24-30.
- Beith, B. (1999, junio 1). *Human Factors and the Future of Telemedicine*. Medical Device & Diagnostic Industry Magazine. <https://www.mddionline.com/news/human-factors-and-future-telemedicine>
- Benitez Sandoval, E., Brown, S., & Velonaki, M. (2018). How the Inclusion of Design Principles Contribute to the Development of Social Robots. *Proceedings of the 30th Australian Computer-Human Interaction Conference (OzCHI'18)*. <https://doi.org/10.1145/3292147.3292239>
- Benyon, D. (2019). Designing User Experience: a guide to HCI. *UX and interaction design*.
- Berg, J. M., Dutton, J. E., & Wrzesniewski, A. (2008). What is job crafting and why does it matter. Retrieved from the website of Positive Organizational Scholarship on April, 15, 2011.
- Beschi, M., Faroni, M., Copot, C., & Pedrocchi, N. (2020). How motion planning affects human factors in human-robot collaboration. *IFAC-PapersOnLine*, 53(5), 744-749. <https://doi.org/10.1016/J.IFACOL.2021.04.167>
- Bhalaji, R. K. A., Bathrinath, S., Ponnambalam, S. G., & Saravanasankar, S. (2021). Analyze the factors influencing human-robot interaction using MCDM method. *Materials Today: Proceedings*, 39, 100-104.
- Bian, Y., Yang, C., Gao, F., Li, H., Zhou, S., Li, H., Sun, X., & Meng, X. (2016). A framework for physiological indicators of flow in VR games: construction and preliminary evaluation. *Personal and Ubiquitous Computing*, 20(5), 821-832. <https://doi.org/https://doi.org/10.1007/s00779-016-0953-5>
- Bitbrain. (2020). *Manual de usuario SennsMetrics 1.0*.
- Bitbrain. (2021). *User Manual: SennsMetrics*.
- Bitbrain Technologies. (2018). *Human Behaviour Research | Software de sincronización biométrica | Bitbrain*. <https://www.bitbrain.com/es/productos-neurotecnologia/software/sennslab>

- Blandino, G. (2023). How to Measure Stress in Smart and Intelligent Manufacturing Systems: A Systematic Review. *Systems* 2023, Vol. 11, Page 167, 11(4), 167. <https://doi.org/10.3390/SYSTEMS11040167>
- Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., Newman, P., Parry, V., Pegman, G., Rodden, T., Sorrell, T., Wallis, M., Whitby, B., & Winfield, A. (2017). Principles of robotics: regulating robots in the real world. *Connection Science*, 29(2), 124-129. <https://doi.org/10.1080/09540091.2016.1271400>
- Bröhl, C., Nelles, J., Brandl, C., Mertens, A., & Nitsch, V. (2019). Human–Robot Collaboration Acceptance Model: Development and Comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics*, 11(5), 709-726. <https://doi.org/10.1007/s12369-019-00593-0>
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7.
- Bruemmer, D. J., Few, D. A., Boring, R. L., Marble, J. L., Walton, M. C., & Nielsen, C. W. (2005). Shared understanding for collaborative control. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans.*, 35(4), 494-504. <https://doi.org/10.1109/TSMCA.2005.850599>
- Bruun, A. (2018). It's not complicated: A study of non-specialists analyzing GSR sensor data to detect UX related events. *ACM International Conference Proceeding Series*, 170-183. <https://doi.org/https://doi.org/10.1145/3240167.3240183>
- Buffet, M. A., Gervais, R. L., Liddle, M., & Eeckelaert, L. (2013). Well-being at work: creating a positive work environment. Literature Review. European Agency for Safety and Health at Work, EU-OSHA. Publications Office of the European Union, Luxembourg. file. *C:/Users/K% C3% A4ytt% C3% A4j% C3% A4/Downloads/well-being-at-work-creating-a-positivework-environment, 20(4)*.
- Butler, J. T., & Agah, A. (2001). Psychological Effects of Behavior Patterns of a Mobile Personal Robot. *Autonomous Robots* 2001 10:2, 10(2), 185-202. <https://doi.org/10.1023/A:1008986004181>
- Cabibihan, J. J., & Chauhan, S. S. (2017). Physiological Responses to Affective Tele-Touch during Induced Emotional Stimuli. *IEEE Transactions on Affective Computing*, 8(1), 108-118. <https://doi.org/https://doi.org/10.1109/TAFFC.2015.2509985>
- Cardoso, A., Colim, A., Bicho, E., Braga, A. C., Menozzi, M., & Arezes, P. (2021). Ergonomics and human factors as a requirement to implement safer collaborative robotic workstations: A literature review. *Safety*, 7(4), 71.
- Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M., & Pecchia, L. (2015). Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. *Biomedical Signal Processing and Control*, 18, 370-377. <https://doi.org/10.1016/J.BSPC.2015.02.012>
- Cazañas, A., de San Miguel, A., & Parra, E. (2017). Estimación del tamaño de la muestra para pruebas de usabilidad. *Enfoque UTE*, 8(1), 172-185. <https://doi.org/http://dx.doi.org/10.29019/enfoqueute.v8n1.126>
- Chang, D., Gu, Z., Li, F., & Jiang, R. (2019). A user-centric smart product-service system development approach: A case study on medication management for the elderly. *Advanced Engineering Informatics*, 42, 100979. <https://doi.org/10.1016/J.AEI.2019.100979>

- Charalambous, G., Fletcher, S. R., & Webb, P. (2017). The development of a Human Factors Readiness Level tool for implementing industrial human-robot collaboration. *The International Journal of Advanced Manufacturing Technology*, 91(5), 2465-2475. <https://doi.org/10.1007/S00170-016-9876-6>
- Chen, M., Nikolaidis, S., Soh, H., Hsu, D., & Srinivasa, S. (2020). Trust-aware decision making for human-robot collaboration: Model learning and planning. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(2), 1-23. <https://doi.org/10.1145/3359616>
- Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 213-218.
- Cho, F. (2019). Analytic Hierarchy Process for Survey Data [R package ahpsurvey version 0.4.1]. En *Mathematical Modelling* (Vol. 9, Números 3-5). Comprehensive R Archive Network (CRAN). [https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)
- Cincuegrani, S. M., Jordà, S., & Våljamäe, A. (2016). Physiopucks: Increasing user motivation by combining tangible and implicit physiological interaction. *ACM Transactions on Computer-Human Interaction*, 23(1), 1-22. <https://doi.org/10.1145/2838732>
- Clarke, R. (1994). Asimov's Laws of Robotics: Implications for Information Technology. 2. *Computer*, 27(1), 57-66. <https://doi.org/10.1109/2.248881>
- Clarkson, E., & Arkin, R. C. (2007). Applying Heuristic Evaluation to Human-Robot Interaction Systems. *Flairs Conference*, 44-49.
- Cohen, Y., Golan, M., Singer, G., & Faccio, M. (2018). Workstation-Operator Interaction in 4.0 Era: WOI 4.0. *IFAC-PapersOnLine*, 51(11), 399-404. <https://doi.org/10.1016/J.IFACOL.2018.08.327>
- Colceriu, C., Leichtmann, B., Brell-Cokcan, S., Jonas, W., & Nitsch, V. (2022). From task analysis to wireframe design: An approach to user-centered design of a GUI for mobile HRI at assembly workplaces. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 876-883.
- Colim, A., Faria, C., Braga, A. C., Sousa, N., Rocha, L., Carneiro, P., Costa, N., & Arezes, P. (2020). Towards an Ergonomic Assessment Framework for Industrial Assembly Workstations—A Case Study. *Applied Sciences* 2020, Vol. 10, Page 3048, 10(9), 3048. <https://doi.org/10.3390/APP10093048>
- Colim, A., Morgado, R., Carneiro, P., Costa, N., Faria, C., Sousa, N., Rocha, L. A., & Arezes, P. (2021). Lean Manufacturing and Ergonomics Integration: Defining Productivity and Wellbeing Indicators in a Human-Robot Workstation. *Sustainability* 2021, Vol. 13, Page 1931, 13(4), 1931. <https://doi.org/10.3390/SU13041931>
- Comisión Europea, D.-G. for R. and I. (2021). *Industry 5.0 - Towards a sustainable, human-centric and resilient European industry*. European Commission. <https://op.europa.eu/en/publication-detail/-/publication/468a892a-5097-11eb-b59f-01aa75ed71a1/>
- Coronado, E., Kiyokawa, T., Ricardez, G. A. G., Ramirez-Alpizar, I. G., Venture, G., & Yamanobe, N. (2022). Evaluating quality in human-robot interaction: A systematic search and classification of performance and human-centered factors, measures and metrics towards an industry 5.0. *Journal of Manufacturing Systems*, 63, 392-410. <https://doi.org/10.1016/J.JMSY.2022.04.007>

- Coronado, E., Villalobos, J., Bruno, B., & Mastrogiovanni, F. (2017). Gesture-based robot control: Design challenges and evaluation with humans. *2017 IEEE international conference on robotics and automation (ICRA)*, 2761-2767.
- Damholdt, M. F., Nørskov, M., Yamazaki, R., Hakli, R., Hansen, C. V., Vestergaard, C., & Seibt, J. (2015). Attitudinal change in elderly citizens toward social robots: The role of personality traits and beliefs about robot functionality. *Frontiers in Psychology*, 6(NOV), 156949. <https://doi.org/10.3389/FPSYG.2015.01701/BIBTEX>
- Daniel, B., Thomessen, T., & Korondi, P. (2013). Simplified Human-Robot Interaction: Modeling and Evaluation. *Modeling, Identification and Control: A Norwegian Research Bulletin*, 34(4), 199-211. <https://doi.org/10.4173/MIC.2013.4.4>
- Danielsson, O., Syberfeldt, A., Brewster, R., & Wang, L. (2017). Assessing Instructions in Augmented Reality for Human-robot Collaborative Assembly by Using Demonstrators. *Procedia CIRP*, 63, 89-94. <https://doi.org/10.1016/J.PROCIR.2017.02.038>
- Dautenhahn, K. (2013). *Human-Robot Interaction | The Encyclopedia of Human-Computer Interaction, 2nd Ed.* <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-robot-interaction>
- Dautenhahn, K., Walters, M., Woods, S., Koay, K. L., Nehaniv, C. L., Sisbot, A., Alami, R., & Siméon, T. (2006). *How may I serve you?* 172. <https://doi.org/10.1145/1121241.1121272>
- David, G. C. (2005). Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occupational medicine*, 55(3), 190-199.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, 13(3), 319-339. <https://doi.org/10.2307/249008>
- Davis, F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International journal of man-machine studies*, 38(3), 475-487.
- Dehais, F., Sisbot, E. A., Alami, R., & Causse, M. (2011). Physiological and subjective evaluation of a human-robot object hand-over task. *Applied Ergonomics*, 42(6), 785-791. <https://doi.org/10.1016/J.APERGO.2010.12.005>
- Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and Human-Robot Co-working. *Procedia Computer Science*, 158, 688-695. <https://doi.org/10.1016/J.PROCS.2019.09.104>
- Devaraj, U. S., Easley, R. F., & Michael Crant, J. (2008). Research Note—How Does Personality Matter? Relating the Five-Factor Model to Technology Acceptance and Use. <https://doi.org/10.1287/isre.1070.0153>, 19(1), 93-105. <https://doi.org/10.1287/ISRE.1070.0153>
- Díaz, C. (2021). SheScience. Science from a gender perspective. *Metode Science Studies Journal*, 7, 94-95.
- Diefenbach, S., Kolb, N., & Hassenzahl, M. (2014). The «hedonic» in human-computer interaction - History, contributions, and future research directions. *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS*, 305-314. <https://doi.org/10.1145/2598510.2598549>
- Dragan, A. D., Bauman, S., Forlizzi, J., & Srinivasa, S. S. (2015). Effects of Robot Motion on Human-Robot Collaboration. *International Conference on Human-Robot Interaction*, 51-58. <https://doi.org/10.1145/2696454.2696473>

- Drury, J. L., Hestand, D., Yanco, H. A., & Scholtz, J. (2004). Design guidelines for improved human-robot interaction. *CHI'04 extended abstracts on Human factors in computing systems*, 1540.
- EFRA (European Factories of the Future Research Association). (s. f.). *Priorities in relation to the Strategic Plan and the Horizon Europe Work Programme* 25-27.
- Empatica. (2022). *E4 wristband | Real-time physiological signals | Wearable PPG, EDA, Temperature, Motion sensors*. <https://www.empatica.com/en-eu/research/e4/>
- Faccio, M., Granata, I., Menini, A., Milanese, M., Rossato, C., Bottin, M., Minto, R., Pluchino, P., Gamberini, L., & Boschetti, G. (2022). Human factors in cobot era: a review of modern production systems features. *Journal of Intelligent Manufacturing*, 1-22.
- Fantini, P., Pinzone, M., & Taisch, M. (2020). Placing the operator at the centre of Industry 4.0 design: Modelling and assessing human activities within cyber-physical systems. *Computers & Industrial Engineering*, 139, 105058.
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, and Computers*, 35(3), 379-383. <https://doi.org/10.3758/BF03195514>
- Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, 22(5), 323-327. <https://doi.org/10.1016/j.intcom.2010.04.004>
- Fishbein, M., & Ajzen, I. (1977). *Belief, attitude, intention, and behavior: An introduction to theory and research*.
- Foglia, P., Zanda, M., & Trading, I. (2014). Towards relating physiological signals to usability metrics: a case study with a web avatar. *WSEAS Transactions on Computers*, 13, 624.
- Frijns, H. A., & Schmidbauer, C. (2021). Design Guidelines for Collaborative Industrial Robot User Interfaces. *IFIP Conference on Human-Computer Interaction*, 407-427.
- Fulfagar, L., Gupta, A., Mathur, A., & Shrivastava, A. (2021). Development and Evaluation of Usability Heuristics for Voice User Interfaces. *International Conference on Research into Design. Smart Innovation, Systems and Technologies*, 221, 375-385. https://doi.org/10.1007/978-981-16-0041-8_32
- Funk, M., Rosen, P. H., & Wischniewski, S. (2020). Human-centered HRI Design—the More Individual the Better? *Behavioral patterns and interaction modelling for personalized human-robot interaction*.
- Galin, R., & Meshcheryakov, R. (2019). Review on human–robot interaction during collaboration in a shared workspace. *International Conference on Interactive Collaborative Robotics*, 63-74.
- Garg, A., Moore, J. S., & Kapellusch, J. M. (2017). The Revised Strain Index: an improved upper extremity exposure assessment model. *Ergonomics*, 60(7), 912-922.
- Gervasi, R., Aliev, K., Mastrogiacomo, L., & Franceschini, F. (2022). User Experience and Physiological Response in Human-Robot Collaboration: A Preliminary Investigation. *Journal of Intelligent & Robotic Systems* 2022 106:2, 106(2), 1-30. <https://doi.org/10.1007/S10846-022-01744-8>
- Gervasi, R., Capponi, M., Mastrogiacomo, L., & Franceschini, F. (2023). Manual assembly and Human–Robot Collaboration in repetitive assembly processes: a structured comparison based on human-centered performances. *The International Journal of Advanced Manufacturing Technology*. <https://doi.org/10.1007/S00170-023-11197-4>

- Gervasi, R., Digiario, F., Mastrogiacomo, L., Maisano, D., & Franceschini, F. (2020). Comparing quality profiles in Human-Robot Collaboration: empirical evidence in the automotive sector. *Proceedings book of the 4th international conference on quality engineering and management. University of Minho, Portugal*, 89-114.
- Gervasi, R., Mastrogiacomo, L., & Franceschini, F. (2020). A conceptual framework to evaluate human-robot collaboration. *International Journal of Advanced Manufacturing Technology*, 108(3), 841-865. <https://doi.org/10.1007/S00170-020-05363-1/FIGURES/11>
- Giakoumis, D., Tzovaras, D., Moustakas, K., & Hassapis, G. (2011). Automatic recognition of boredom in video games using novel biosignal moment-based features. *IEEE Transactions on Affective Computing*, 2(3), 119-133. <https://doi.org/https://doi.org/10.1109/T-AFFC.2011.4>
- Goodrich, M. A., & Schultz, A. C. (2008). *Human-robot interaction: a survey*. Now Publishers Inc.
- Gopinath, V., Ore, F., & Johansen, K. (2017). Safe Assembly Cell Layout through Risk Assessment – An Application with Hand Guided Industrial Robot. *Procedia CIRP*, 63, 430-435. <https://doi.org/10.1016/J.PROCIR.2017.03.160>
- Gopsill, J. A., McAlpine, H. C., & Hicks, B. J. (2015). Supporting engineering design communication using a custom-built social media tool – PartBook. *Advanced Engineering Informatics*, 29(3), 523-548. <https://doi.org/10.1016/J.AEI.2015.04.008>
- Gramann, K., Onton, J., Riccobon, D., Mueller, H. J., Bardins, S., & Makeig, S. (2010). Human brain dynamics accompanying use of egocentric and allocentric reference frames during navigation. *Journal of Cognitive Neuroscience*, 22(12), 2836-2849. <https://doi.org/https://doi.org/10.1162/jocn.2009.21369>
- Grandi, F., Peruzzini, M., Cavallaro, S., Prati, E., & Pellicciari, M. (2022). Creation of a UX index to design human tasks and workstations. *International Journal of Computer Integrated Manufacturing*, 35(1), 4-20. <https://doi.org/10.1080/0951192X.2021.1972470>
- Gualtieri, L., Monizza, G. P., Rauch, E., Vidoni, R., & Matt, D. T. (2020). From Design for Assembly to Design for Collaborative Assembly - Product Design Principles for Enhancing Safety, Ergonomics and Efficiency in Human-Robot Collaboration. *Procedia CIRP*, 91, 546-552. <https://doi.org/10.1016/J.PROCIR.2020.02.212>
- Gualtieri, L., Palomba, I., Merati, F. A., Rauch, E., & Vidoni, R. (2020). Design of Human-Centered Collaborative Assembly Workstations for the Improvement of Operators' Physical Ergonomics and Production Efficiency: A Case Study. *Sustainability 2020, Vol. 12, Page 3606, 12(9)*, 3606. <https://doi.org/10.3390/SU12093606>
- Gualtieri, L., Rauch, E., & Vidoni, R. (2021). Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, 67, 101998. <https://doi.org/10.1016/J.RCIM.2020.101998>
- Gualtieri, L., Rauch, E., & Vidoni, R. (2022). Development and validation of guidelines for safety in human-robot collaborative assembly systems. *Computers & Industrial Engineering*, 163, 107801. <https://doi.org/10.1016/j.cie.2021.107801>
- Gualtieri, L., Rauch, E., Vidoni, R., & Matt, D. T. (2020). Safety, Ergonomics and Efficiency in Human-Robot Collaborative Assembly: Design Guidelines and Requirements. *Procedia CIRP*, 91, 367-372. <https://doi.org/10.1016/J.PROCIR.2020.02.188>

- Guerin, K. R., Lea, C., Paxton, C., & Hager, G. D. (2015). A framework for end-user instruction of a robot assistant for manufacturing. *2015 IEEE international conference on robotics and automation (ICRA)*, 6167-6174.
- Gulliver, S. R., & Ghinea, G. (2006). Defining user perception of distributed multimedia quality. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2(4), 241-257. <https://doi.org/https://doi.org/10.1145/1201730.1201731>
- Gupta, K., Lee, G. A., & Billinghurst, M. (2016). Do you see what i see? the effect of gaze tracking on task space remote collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 22(11), 2413-2422. <https://doi.org/https://doi.org/10.1109/TVCG.2016.2593778>
- Hadorn, B., Courant, M., & Hirsbrunner, B. (2015). Holistic integration of enactive entities into cyber physical systems. *2015 IEEE 2nd international conference on cybernetics (CYBCONF)*, 281-286.
- Hanajima, N., Goto, T., Ohta, Y., Hikita, H., & Yamashita, M. (2017). Effects of movement speed and predictability in human-robot collaboration. *Human Factors and Ergonomics in Manufacturing*, 27(4), 197-209. <https://doi.org/10.1002/hfm.20703>
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., Visser, E. J. de, & Parasuraman, R. (2011). A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction: *Human Factors*, 53(5), 517-527. <https://doi.org/10.1177/0018720811417254>
- Hancock, P. A., Pepe, A. A., & Murphy, L. L. (2005). Hedonomics: The power of positive and pleasurable ergonomics. *Ergonomics in Design*, 13(1), 8-14. https://doi.org/10.1177/106480460501300104/ASSET/106480460501300104.FP.PNG_V03
- Harriott, C. E., Zhang, T., & Adams, J. A. (2013). Assessing physical workload for human-robot peer-based teams. *International Journal of Human-Computer Studies*, 71(7-8), 821-837. <https://doi.org/10.1016/J.IJHCS.2013.04.005>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, 52(C), 139-183. [https://doi.org/https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/https://doi.org/10.1016/S0166-4115(08)62386-9)
- Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, 13(4), 481-499.
- Hassenzahl, M. (2007). The hedonic/pragmatic model of user experience. *Towards a UX manifesto*, 10, 2007.
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. En *Mensch & computer 2003* (pp. 187-196). Springer.
- Helander, M. G. (1997). The human factors profession (Chapter 1). En *Handbook of Human Factors and Ergonomics*.
- Hentout, A., Aouache, M., Maoudj, A., & Akli, I. (2019). Human-robot interaction in industrial collaborative robotics: a literature review of the decade 2008-2017. *Advanced Robotics*, 33(15-16), 764-799. <https://doi.org/https://doi.org/10.1080/01691864.2019.1636714>
- Heyer, C. (2010). Human-robot interaction and future industrial robotics applications. *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4749-4754.

- Hietanen, A., Pieters, R., Lanz, M., Latokartano, J., & Kämäräinen, J. K. (2020). AR-based interaction for human-robot collaborative manufacturing. *Robotics and Computer-Integrated Manufacturing*, 63, 101891. <https://doi.org/10.1016/J.RCIM.2019.101891>
- Hignett, S., & McAtamney, L. (2000). Rapid entire body assessment (REBA). *Applied ergonomics*, 31(2), 201-205.
- Hinderks, A., Schrepp, M., & Thomaschewski, J. (2018). *User Experience Questionnaire (UEQ)*. <https://www.ueq-online.org/>
- Ho, D., Newell, G., & Walker, A. (2005). The importance of property-specific attributes in assessing CBD office building quality. *Journal of Property Investment and Finance*, 23(5), 424-444. <https://doi.org/10.1108/14635780510616025>
- Holz, E. M., Höhne, J., Staiger-Sälzer, P., Tangermann, M., & Kübler, A. (2013). Brain-computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users. *Artificial Intelligence in Medicine*, 59(2), 111-120. <https://doi.org/https://doi.org/10.1016/j.artmed.2013.08.001>
- Horizon Europe guidance on gender equality plans - Publications Office of the EU. (s. f.). Recuperado 10 de noviembre de 2021, de <https://op.europa.eu/en/publication-detail/-/publication/ffc06c3-200a-11ec-bd8e-01aa75ed71a1>
- Horváth, I. (2007). Comparison of three methodological approaches of design research. *DS 42: Proceedings of ICED 2007, the 16th International Conference on Engineering Design, Paris, France, 28.-31.07. 2007*, 361-362.
- Huynh, S., Kim, S., Ko, J., Balan, R. K., & Lee, Y. (2018). EngageMon. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1), 1-27. <https://doi.org/https://doi.org/10.1145/3191745>
- International Ergonomics Association. (2019). *Definition and Domains of Ergonomics*. Ergonomics. <https://doi.org/10.1080/00140139.2012.661087>
- Irfan, B., Ramachandran, A., Spaulding, S., Glas, D. F., Leite, I., & Koay, K. L. (2019). Personalization in long-term human-robot interaction. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 685-686.
- ISO 6385. (2016). *Principios ergonómicos para el diseño de sistemas de trabajo*. <https://tienda.aenor.com/norma-une-en-iso-6385-2016-n0057285>
- ISO 8373. (2021). *Robotics — Vocabulary*. <https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-3:v1:en>
- ISO 9241-11. (2018). *Ergonomics of human-system interaction — Part...* <https://www.une.org/encuentra-tu-norma/busca-tu-norma/iso?c=063500>
- ISO 9241-210. (2019). Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. En *International Standard: Vol. 2.Udgave*.
- ISO 10218. (2011). *ISO 10218: Robots and Robotic Devices: Safety Requirements for Industrial Robots*. International Organization for Standardization.
- ISO/TS 15066. (2016). *Robots and robotic devices—collaborative robots*.
- Jahangiri, P., Saberi, M. K., & Vakilimofrad, H. (2021). Development and psychometric evaluation of the cloud computing acceptance questionnaire for academic libraries. *The Journal of Academic Librarianship*, 47(5), 102395. <https://doi.org/10.1016/J.ACALIB.2021.102395>

- Jiang, S., Li, Z., Zhou, P., & Li, M. (2019). Memento: An emotion-driven lifelogging system with wearables. *ACM Transactions on Sensor Networks*, 15(1), 1-23. <https://doi.org/https://doi.org/10.1145/3281630>
- Joose, M., Lohse, M., Berkel, N. van, Sardar, A., & Evers, V. (2021). Making Appearances: How Robots Should Approach People. *ACM Transactions on Human-Robot Interaction (THRI)*, 10(1). <https://doi.org/10.1145/3385121>
- Joose, M., Sardar, A., Lohse, M., Evers, V., Joosse, M., Lohse, M., Evers, V., Lohse, M., Evers, V., & Sardar, A. (2013). BEHAVE-II: The Revised Set of Measures to Assess Users' Attitudinal and Behavioral Responses to a Social Robot. *Int J Soc Robot*, 5, 379-388. <https://doi.org/10.1007/s12369-013-0191-1>
- Joyce, A. (2019, octubre 20). *How to Measure Learnability of a User Interface*. <https://www.nngroup.com/articles/measure-learnability/>
- Kahn Jr, P. H., Ishiguro, H., Friedman, B., Kanda, T., Freier, N. G., Severson, R. L., & Miller, J. (2007). What is a Human?: Toward psychological benchmarks in the field of human-robot interaction. *Interaction Studies*, 8(3), 363-390. <https://doi.org/https://doi.org/10.1075/is.8.3.04kah>
- Kaiser, L., Schlotzhauer, A., & Brandstötter, M. (2018). Safety-related risks and opportunities of key design-aspects for industrial human-robot collaboration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11097 LNAI, 95-104. https://doi.org/10.1007/978-3-319-99582-3_11
- Karabegović, I., Turmanidze, R., & Dašić, P. (2020). Global trend of implementation of industrial robots relating to industry 4.0. *Design, Simulation, Manufacturing: The Innovation Exchange*, 147-155.
- Karabegović, I., Turmanidze, R., & Dašić, P. (2022). Structural Network for the Implementation of "Industry 4.0" in Production Processes. *Industry 4.0*, 7(1), 3-6.
- Karegar, F., Pettersson, J. S., & Fischer-Hübner, S. (2020). The dilemma of user engagement in privacy notices: Effects of interaction modes and habituation on user attention. *ACM Transactions on Privacy and Security*, 23(1), 1-38. <https://doi.org/https://doi.org/10.1145/3372296>
- Katsigiannis, S., Willis, R., & Ramzan, N. (2019). A QoE and Simulator Sickness Evaluation of a Smart-Exercise-Bike Virtual Reality System via User Feedback and Physiological Signals. *IEEE Transactions on Consumer Electronics*, 65(1), 119-127. <https://doi.org/https://doi.org/10.1109/TCE.2018.2879065>
- Keebler, J. R., Jentsch, F., Fincannon, T., & Hudson, I. (2012). Applying team heuristics to future human-robot systems. *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 169-170.
- Kerr, J., & Coviello, N. (2019). Formation and Constitution of Effectual Networks: A Systematic Review and Synthesis. *International Journal of Management Reviews*, 21(3), 370-397. <https://doi.org/https://doi.org/10.1111/ijmr.12194>
- Khalid, A., Kirisci, P., Ghrairi, Z., Pannek, J., & Thoben, K.-D. (2017). Safety Requirements in Collaborative Human-Robot Cyber-Physical System. *Lecture Notes in Logistics*, 41-51. https://doi.org/10.1007/978-3-319-45117-6_4
- Khamaisi, R. K., Brunzini, A., Grandi, F., Peruzzini, M., & Pellicciari, M. (2022). UX assessment strategy to identify potential stressful conditions for workers. *Robotics and Computer-Integrated Manufacturing*, 78, 102403. <https://doi.org/10.1016/J.RCIM.2022.102403>

- Khamaisi, R. K., Grandi, F., Prati, E., Peruzzini, M., & Pellicciari, M. (2022). A comprehensive UX index to evaluate industrial tasks from a human-centered perspective. *IEEE International Workshop on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering, MetroXRINE 2022 - Proceedings*, 52-57. <https://doi.org/10.1109/METROXRINE54828.2022.9967677>
- Kildal, J., Martín, M., Ipiña, I., & Murtua, I. (2019). Empowering assembly workers with cognitive disabilities by working with collaborative robots: a study to capture design requirements. *Procedia CIRP*, 81, 797-802. <https://doi.org/10.1016/J.PROCIR.2019.03.202>
- Kim, Y., & Mutlu, B. (2014). How social distance shapes human–robot interaction. *International Journal of Human-Computer Studies*, 72(12), 783-795.
- Kirakowski, J., & Cierlik, B. (1998). Measuring the usability of web sites. *Proceedings of the Human Factors and Ergonomics Society annual meeting*, 42(4), 424-428. <https://doi.org/10.1177/154193129804200405>
- Kirakowski, J., & Corbett, M. (1993). SUMI: The software usability measurement inventory. *British journal of educational technology*, 24(3), 210-212.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*.
- Kokotinis, G., Michalos, G., Arkouli, Z., & Makris, S. (2023). On the quantification of human-robot collaboration quality. *International Journal of Computer Integrated Manufacturing*, 1-18. <https://doi.org/10.1080/0951192X.2023.2189304>
- Koppenborg, M., Nickel, P., Naber, B., Lungfiel, A., & Huelke, M. (2017). Effects of movement speed and predictability in human–robot collaboration. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 27(4), 197-209. <https://doi.org/10.1002/HFM.20703>
- Kuijper, P. P. F. M., Visser, B., & Kemper, H. C. G. (1999). Job rotation as a factor in reducing physical workload at a refuse collecting department. *Ergonomics*, 42(9), 1167-1178.
- Lachner, F., Naegelein, P., Kowalski, R., Spann, M., & Butz, A. (2016). Quantified UX: Towards a common organizational understanding of user experience. *ACM International Conference Proceeding Series, 23-27-October-2016*. <https://doi.org/10.1145/2971485.2971501>
- Laros, F. J. M., & Steenkamp, J.-B. E. M. (2005). Emotions in consumer behavior: a hierarchical approach. *Journal of business Research*, 58(10), 1437-1445.
- Lasota, P. A., & Shah, J. A. (2015). Analyzing the Effects of Human-Aware Motion Planning on Close-Proximity Human–Robot Collaboration. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(1), 21-33. <https://doi.org/10.1177/0018720814565188>
- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. *Symposium of the Austrian HCI and usability engineering group*, 63-76.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4), 563-575.
- LBR iiwa | KUKA AG. (s.f.). Recuperado 2 de noviembre de 2021, de <https://www.kuka.com/en-de/products/robot-systems/industrial-robots/lbr-iiwa>
- Lee, K. C., & Chung, N. (2009). Understanding factors affecting trust in and satisfaction with mobile banking in Korea: A modified DeLone and McLean’s model perspective. *Interacting with computers*, 21(5-6), 385-392.

- Lee, Y. C., Lin, W. C., Cherng, F. Y., & Ko, L. W. (2016). A Visual Attention Monitor Based on Steady-State Visual Evoked Potential. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(3), 399-408. <https://doi.org/https://doi.org/10.1109/TNSRE.2015.2501378>
- Lewis, J. R. (1991a). An after-scenario questionnaire for usability studies: psychometric evaluation over three trials. *ACM SIGCHI Bulletin*, 23(4), 79.
- Lewis, J. R. (1991b). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM Sigchi Bulletin*, 23(1), 78-81.
- Lewis, J. R. (1992a). Psychometric evaluation of the computer system usability questionnaire: The CSUQ (Tech. Report 54.723). Boca Raton, FL: International Business Machines Corporation.
- Lewis, J. R. (1992b). Psychometric evaluation of the post-study system usability questionnaire: The PSSUQ. *Proceedings of the human factors society annual meeting*, 36(16), 1259-1260.
- Lewis, J. R. (1994). Sample sizes for usability studies: Additional considerations. *Human factors*, 36(2), 368-378. <https://doi.org/https://doi.org/10.1177/001872089403600215>
- Lewis, J. R. (2009). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. <https://doi.org/10.1080/10447319509526110>, 7(1), 57-78. <https://doi.org/10.1080/10447319509526110>
- Lewis, J. R., & Mayes, D. K. (2014). Development and psychometric evaluation of the emotional metric outcomes (EMO) questionnaire. *International Journal of Human-Computer Interaction*, 30(9), 685-702.
- Lewis, J. R., Utesch, B. S., & Maher, D. E. (2013). UMUX-LITE-When There's No Time for the SUS. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2470654>
- Lewis, J. R., Utesch, B. S., & Maher, D. E. (2015). Measuring perceived usability: The SUS, UMUX-LITE, and AltUsability. *International Journal of Human-Computer Interaction*, 31(8), 496-505.
- Lewis, T., Drury, J. L., & Brandon, B. (2014, noviembre). Evaluating Mobile Remote Presence (MRP) Robots. *Proceedings of the 18th International Conference on Supporting Group Work*.
- Lin, H. X., Choong, Y.-Y., & Salvendy, G. (1997). A proposed index of usability: a method for comparing the relative usability of different software systems. *Behaviour & information technology*, 16(4-5), 267-277.
- Lindblom, J., & Alenljung, B. (2020). The ANEMONE: Theoretical Foundations for UX Evaluation of Action and Intention Recognition in Human-Robot Interaction. *Sensors*, 20(15), 4284. <https://doi.org/https://doi.org/10.3390/s20154284>
- Liu, Y., Ritchie, J. M., Lim, T., Kosmadoudi, Z., Sivanathan, A., & Sung, R. C. W. (2014). A fuzzy psycho-physiological approach to enable the understanding of an engineer's affect status during CAD activities. *CAD Computer Aided Design*, 54, 19-38. <https://doi.org/https://doi.org/10.1016/j.cad.2013.10.007>
- López-Gil, J. M., Navarro-Molina, C., García, R., & Aleixandre-Benavent, R. (2010). Análisis de la arquitectura de webs mediante tests de estrés de navegación, de usabilidad y eye tracking. *Profesional de la Información*, 19(4), 359-367. <https://doi.org/https://doi.org/10.3145/epi.2010.jul.04>

- Lund, A. M. (2001). Measuring usability with the use questionnaire¹². *Usability interface*, 8(2), 3-6.
- Luther, L., Tiberius, V., & Brem, A. (2020). User Experience (UX) in business, management, and psychology: A bibliometric mapping of the current state of research. *Multimodal Technologies and Interaction*, 4(2), 18.
- Maddikunta, P. K. R., Pham, Q. V., B, P., Deepa, N., Dev, K., Gadekallu, T. R., Ruby, R., & Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, 26, 100257. <https://doi.org/10.1016/J.JII.2021.100257>
- Maguire, M. (2019). Development of a heuristic evaluation tool for voice user interfaces. *International conference on human-computer interaction*, 212-225.
- Maia, C. L. B., & Furtado, E. S. (2016a). A study about psychophysiological measures in user experience monitoring and evaluation. *Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems*, 1-9. <https://doi.org/https://doi.org/10.1145/3033701.3033708>
- Maia, C. L. B., & Furtado, E. S. (2016b). A systematic review about user experience evaluation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9746, 445-455. https://doi.org/https://doi.org/10.1007/978-3-319-40409-7_42
- Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology*, 25(2), 141-158. <https://doi.org/10.1080/01449290500331156>
- Mangion, R. S., Garg, L., Garg, G., & Falzon, O. (2020). Emotional Testing on Facebook's User Experience. *IEEE Access*, 8, 58250-58259. <https://doi.org/https://doi.org/10.1109/ACCESS.2020.2981418>
- Manning, M. D., Harriott, C. E., Hayes, S. T., Adams, J. A., & Seiffert, A. E. (2015). Heuristic evaluation of swarm metrics' effectiveness. *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction extended abstracts*, 17-18.
- Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European Portuguese Validation of the System Usability Scale (SUS). *Procedia Computer Science*, 67, 293-300. <https://doi.org/10.1016/J.PROCS.2015.09.273>
- Marvel, J. A., Bagchi, S., Zimmerman, M., & Antonishek, B. (2020). Towards effective interface designs for collaborative HRI in manufacturing: metrics and measures. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(4), 1-55. <https://doi.org/https://doi.org/10.1145/3385009>
- Maslow, A. H. (1943). A Theory of Human Motivation. *Psychological Review*, 50, 370-396. <http://psychclassics.yorku.ca/Maslow/motivation.htm>
- Matheson, E., Minto, R., Zampieri, E. G. G., Faccio, M., & Rosati, G. (2019). Human-Robot Collaboration in Manufacturing Applications: A Review. *Robotics 2019, Vol. 8, Page 100*, 8(4), 100. <https://doi.org/10.3390/ROBOTICS8040100>
- Maurice, P., Padois, V., Measson, Y., & Bidaud, P. (2017). Human-oriented design of collaborative robots. *International Journal of Industrial Ergonomics*, 57, 88-102. <https://doi.org/10.1016/J.ERGON.2016.11.011>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review*, 20(3), 709-734.

- Mazmela, M. (2020). *Nuevo modelo para evaluar el nivel de aceptación de las tecnologías desde la perspectiva de la interacción y percepción de usuario* [Mondragon Goi Eskola Politeknikoja]. <https://hdl.handle.net/20.500.11984/5146>
- Mazmela, M., Lasa, G., & Apraiz, A. (2020). Aproximación multimétodo para medir el impacto de los factores de diseño en la apropiación de un software de inteligencia competitiva. *DYNA new technologies, ISSN-e 2386-8406, Vol. 7, N.º. 1, 2020, 7(1), 8*. <https://dialnet.unirioja.es/servlet/articulo?codigo=7702107&info=resumen&idioma=SPA>
- Mazmela, M., Lasa, G., Aranburu, E., Tomás, P., & Anaya, M. (2019). *La evaluación de la ejecución de tareas por parte de los usuarios en interfaces industriales mediante el cuestionario USE*.
- McAtamney, L., & Corlett, E. N. (1993). RULA: a survey method for the investigation of work-related upper limb disorders. *Applied ergonomics, 24(2)*, 91-99.
- McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & Kaliouby, R. el. (2016). AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit. *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems, 3723-3726*.
- Meissner, A., Trübswetter, A., Conti-Kufner, A. S., & Schmidtler, J. (2020). Friend or Foe? Understanding Assembly Workers' Acceptance of Human-robot Collaboration. *ACM Transactions on Human-Robot Interaction (THRI), 10(1)*. <https://doi.org/10.1145/3399433>
- Menges, R., Kumar, C., & Staab, S. (2019). Improving user experience of eye tracking-based interaction: Introspecting and adapting interfaces. *ACM Transactions on Computer-Human Interaction, 26(6)*. <https://doi.org/https://doi.org/10.1145/3338844>
- Meza-Kubo, V., Morán, A. L., Carrillo, I., Galindo, G., & García-Canseco, E. (2016). Assessing the user experience of older adults using a neural network trained to recognize emotions from brain signals. *Journal of Biomedical Informatics, 62*, 202-209. <https://doi.org/https://doi.org/10.1016/j.jbi.2016.07.004>
- Minge, M., Thüring, M., Wagner, I., & Kuhr, C. v. (2017). The meCUE questionnaire: a modular tool for measuring user experience. En *Advances in Ergonomics Modeling, Usability & Special Populations* (pp. 115-128). Springer.
- Molich, R., & Nielsen, J. (1990). Improving a human-computer dialogue. *Communications of the ACM, 33(3)*, 338-348.
- Mondragon Unibertsitatea. (2020, junio 16). *El nuevo laboratorio industria 4.0 de Mondragon Unibertsitatea abre sus puertas a empresas de ciberseguridad para colaborar en investigación*. <https://www.mondragon.edu/es/-/industria-4-laborategi-berria>
- Moroney, W. F., & Lilienthal, M. G. (2008). Human factors in simulation and training. *Human Factors in Simulation and Training. CRC Press*, 3-38.
- Morrow, P. B., & Fiore, S. M. (2012). *Supporting Human-Robot Teams in Social Dynamicism: an overview of the Metaphoric Inference Framework; Supporting Human-Robot Teams in Social Dynamicism: an overview of the Metaphoric Inference Framework*. <https://doi.org/10.1177/1071181312561344>
- Moshagen, M., & Thielsch, M. (2013). A short version of the visual aesthetics of websites inventory. *Behaviour & Information Technology, 32(12)*, 1305-1311.
- Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies, 68(10)*, 689-709. <https://doi.org/10.1016/J.IJHCS.2010.05.006>

- Müller, S. L., & Richert, A. (2018). The big-five personality dimensions and attitudes to-wards robots: A cross sectional study. *ACM International Conference Proceeding Series*, 405-408. <https://doi.org/10.1145/3197768.3203178>
- Murphy, R. R., & Tadokoro, S. (2019). User Interfaces for Human-Robot Interaction in Field Robotics. *Springer Tracts in Advanced Robotics*, 128, 507-528. https://doi.org/10.1007/978-3-030-05321-5_11
- Nahavandi, S. (2019). Industry 5.0—A human-centric solution. *Sustainability*, 11(16), 4371.
- Negi, P. (2021). Exploring Human-Robot Interaction in Mechatronic Systems: Challenges and Opportunities. *Mathematical Statistician and Engineering Applications*, 70(1), 386-394.
- Neumann, D. L., & Westbury, H. R. (2011). The psychophysiological measurement of empathy. *Psychology of empathy*, 119-142.
- Nguyen, H. N., Lasa, G., Iriarte, I., Atxa, A., Unamuno, G., & Galfarsoro, G. (2022). Human-centered design for advanced services: A multidimensional design methodology. *Advanced Engineering Informatics*, 53, 101720.
- Nguyen Ngoc, H., Lasa, G., & Iriarte, I. (2022). Human-centred design in industry 4.0: case study review and opportunities for future research. *Journal of Intelligent Manufacturing*, 33(1), 35-76.
- Nicolò, A., Massaroni, C., Schena, E., & Sacchetti, M. (2020). The Importance of Respiratory Rate Monitoring: From Healthcare to Sport and Exercise. *Sensors*, 20(21), 6396. <https://doi.org/10.3390/S20216396>
- Nidhra, S., Yanamadala, M., Afzal, W., & Torkar, R. (2013). Knowledge transfer challenges and mitigation strategies in global software development—A systematic literature review and industrial validation. *International Journal of Information Management*, 33(2), 333-355. <https://doi.org/10.1016/J.IJINFOMGT.2012.11.004>
- Nielsen, J. (2000). *Why you only need to test with 5 users*. Useit. com Alertbox.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Conference on Human Factors in Computing Systems - Proceedings*, 249-256. <https://doi.org/10.1145/97243.97281>
- Niksirat, K. S., Silpasuwanchai, C., Cheng, P., & Ren, X. (2019). Attention regulation framework: Designing self-regulated mindfulness technologies. *ACM Transactions on Computer-Human Interaction*, 26(6). <https://doi.org/https://doi.org/10.1145/3359593>
- Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2004). *Psychology in Human-Robot Communication: An Attempt through Investigation of Negative Attitudes and Anxiety toward Robots*. 04TH8759.
- Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.
- Norman, D. (2014). *Things that make us smart: Defending human attributes in the age of the machine*. Diversion Books.
- Norman, D. A. (1986). Cognitive engineering. *User centered system design*, 31, 61.
- Norman, D. A. (1991). Cognitive artifacts. *Designing interaction: Psychology at the human-computer interface*, 1(1), 17-38.
- Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Books.
- Norman, D. A. (2005). Human-centered design considered harmful. *interactions*, 12(4), 14-19.

- Nourbakhsh, N., Chen, F., Wang, Y., & Calvo, R. A. (2017). Detecting users' cognitive load by galvanic skin response with affective interference. *ACM Transactions on Interactive Intelligent Systems*, 7(3), 1-20. <https://doi.org/10.1145/2960413>
- Oliveira, I. I. P., & Guimarães, N. M. (2013). Practical Neurophysiological Analysis of Readability as a Usability Dimension. *International Conference on Human Factors in Computing and Informatics*, 194-211.
- Onnasch, L., & Roesler, E. (2021). A taxonomy to structure and analyze human–robot interaction. *International Journal of Social Robotics*, 13(4), 833-849. <https://doi.org/10.1007/s12369-020-00666-5>
- Oron-Gilad, T., & Hancock, P. A. (2017). From Ergonomics to Hedonomics: Trends in Human Factors and Technology—The Role of Hedonomics Revisited. En *Emotions and Affect in Human Factors and Human-Computer Interaction* (pp. 185-194). Academic Press. <https://doi.org/10.1016/B978-0-12-801851-4.00007-0>
- Pacaux-Lemoine, M.-P., Berdal, Q., Guérin, C., Rauffet, P., Chauvin, C., & Trentesaux, D. (2021). Designing human–system cooperation in industry 4.0 with cognitive work analysis: a first evaluation. *Cognition, Technology & Work*, 1-19.
- Paliga, M. (2023). The Relationships of Human-Cobot Interaction Fluency with Job Performance and Job Satisfaction among Cobot Operators—The Moderating Role of Workload. *International Journal of Environmental Research and Public Health*, 20(6), 5111.
- Pandey, A. K., & Gelin, R. (2018). A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3), 40-48.
- Pantano, M., Regulin, D., Lutz, B., & Lee, D. (2020). A human-cyber-physical system approach to lean automation using an industrie 4.0 reference architecture. *Procedia Manufacturing*, 51, 1082-1090. <https://doi.org/10.1016/J.PROMFG.2020.10.152>
- Pardo, D., Möller, L., Neunert, M., Winkler, A. W., & Buchli, J. (2016). Evaluating direct transcription and nonlinear optimization methods for robot motion planning. *IEEE Robotics and Automation Letters*, 1(2), 946-953.
- Peruzzini, M., Grandi, F., & Pellicciari, M. (2017). Benchmarking of Tools for User Experience Analysis in Industry 4.0. *Procedia Manufacturing*, 11, 806-813. <https://doi.org/10.1016/J.PROMFG.2017.07.182>
- Peruzzini, M., Grandi, F., & Pellicciari, M. (2018). How to analyse the workers' experience in integrated product-process design. *Journal of Industrial Information Integration*, 12, 31-46. <https://doi.org/10.1016/j.jii.2018.06.002>
- Peruzzini, M., Grandi, F., & Pellicciari, M. (2020). Exploring the potential of Operator 4.0 interface and monitoring. *Computers and Industrial Engineering*, 139, 105600. <https://doi.org/10.1016/j.cie.2018.12.047>
- Prati, E., Peruzzini, M., Pellicciari, M., & Raffaelli, R. (2021). How to include User eXperience in the design of Human-Robot Interaction. *Robotics and Computer-Integrated Manufacturing*, 68, 102072. <https://doi.org/10.1016/J.RCIM.2020.102072>
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1-15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)
- Prewett, M. S., Johnson, R. C., Saboe, K. N., Elliott, L. R., & Coovert, M. D. (2010). Managing workload in human–robot interaction: A review of empirical studies. *Computers in Human Behavior*, 26(5), 840-856.

- Punnett, L., & Wegman, D. H. (2004). Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *Journal of Electromyography and Kinesiology*, 14(1), 13-23. <https://doi.org/10.1016/J.JELEKIN.2003.09.015>
- Qbilat, M., & Iglesias, A. (2018). Accessibility guidelines for tactile displays in human-robot interaction. A comparative study and proposal. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10897 LNCS, 217-220. https://doi.org/10.1007/978-3-319-94274-2_29/COVER
- Qbilat, M., Iglesias, A., & Belpaeme, T. (2021). A Proposal of Accessibility Guidelines for Human-Robot Interaction. *Electronics 2021, Vol. 10, Page 561, 10(5)*, 561. <https://doi.org/10.3390/ELECTRONICS10050561>
- Quiñones, D., & Rusu, C. (2017). How to develop usability heuristics: A systematic literature review. *Computer Standards & Interfaces*, 53, 89-122. <https://doi.org/10.1016/J.CSI.2017.03.009>
- Rajanen, D., Clemmensen, T., Iivari, N., Inal, Y., Rizvanoğlu, K., Sivaji, A., & Roche, A. (2017). UX professionals' definitions of usability and UX – A comparison between Turkey, Finland, Denmark, France and Malaysia. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10516 LNCS, 218-239. https://doi.org/10.1007/978-3-319-68059-0_14/TABLES/9
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality*, 41(1), 203-212.
- Rana, A., & Kumar Dubey, S. (2010). Analytical Roadmap to Usability Definitions and Decompositions. *International Journal of Engineering Science and Technology*, 2(9), 4723-4729. <https://www.researchgate.net/publication/282848700>
- Randolph, A. B., & Jackson, M. M. M. (2010). Assessing fit of nontraditional assistive technologies. *ACM Transactions on Accessible Computing*, 2(4), 1-31. <https://doi.org/https://doi.org/10.1145/1786774.1786777>
- Rawassizadeh, R., Price, B. A., & Petre, M. (2015). Wearables: Has The age of smartwatches finally arrived? En *Communications of the ACM* (Vol. 58, Número 1, pp. 45-47). Association for Computing Machinery. <https://doi.org/https://doi.org/10.1145/2629633>
- Reich, N., & Eyssel, F. (2013). Attitudes towards service robots in domestic environments: The role of personality characteristics, individual interests, and demographic variables. *Paladyn, Journal of Behavioral Robotics*, 4(2), 123-130. <https://doi.org/10.2478/PJBR-2013-0014>
- Robinson, J., Lanius, C., & Weber, R. (2018). The past, present, and future of UX empirical research. *Communication Design Quarterly*, 5(3), 10-23. <https://doi.org/https://doi.org/10.1145/3188173.3188175>
- Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. En *An integrated approach to communication theory and research* (pp. 432-448). Routledge.
- Romero, D., Mattsson, S., Fast-Berglund, Å., Wuest, T., Gorecky, D., & Stahre, J. (2018). Digitalizing occupational health, safety and productivity for the operator 4.0. *IFIP Advances in Information and Communication Technology*, 536, 473-481. https://doi.org/10.1007/978-3-319-99707-0_59/COVER
- Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Fast-Berglund, Å., & Gorecky, D. (2016). Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial

- revolution technologies. *proceedings of the international conference on computers and industrial engineering (CIE46), Tianjin, China*, 29-31.
- Romero, D., Wuest, T., Stahre, J., & Gorecky, D. (2017). Social factory architecture: social networking services and production scenarios through the social internet of things, services and people for the social operator 4.0. *IFIP international conference on advances in production management systems*, 265-273.
- Roto, V., Law, E., Vermeeren, A., & Hoonhout, J. (2011, febrero 11). User Experience White Paper: Bringing clarity to the concept of user experience. *Outcome of Dagstuhl Seminar 10373: Demarcating User Experience*. www.allaboutux.org/ux-definitions.
- Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. John Wiley & Sons.
- Rücker, D., Hornfeck, R., & Paetzold, K. (2019). Investigating ergonomics in the context of human-robot collaboration as a sociotechnical system. *Advances in Intelligent Systems and Computing*, 784, 127-135. https://doi.org/10.1007/978-3-319-94346-6_12/COVER
- Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, 9(1), 54-89.
- Saaty, T. L. (2004). Fundamentals of the analytic network process — Dependence and feedback in decision-making with a single network. *Journal of Systems Science and Systems Engineering* 2004 13:2, 13(2), 129-157. <https://doi.org/10.1007/S11518-006-0158-Y>
- Salanova, M., Del Líbano, M., Llorens, S., & Schaufeli, W. B. (2014). Engaged, workaholic, burned-out or just 9-to-5? Toward a typology of employee well-being. *Stress and Health*, 30(1), 71-81.
- Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, 141-148.
- Sanches, P., Höök, K., Sas, C., & Ståhl, A. (2019). Ambiguity as a resource to inform proto-practices: The case of skin conductance. *ACM Transactions on Computer-Human Interaction*, 26(4), 1-32. <https://doi.org/https://doi.org/10.1145/3318143>
- Sarker, S., & Wells, J. D. (2003). Understanding mobile handheld device use and adoption. *Communications of the ACM*, 46(12), 35-40.
- Sauer, J., Sonderegger, A., & Schmutz, S. (2020). Usability, user experience and accessibility: towards an integrative model. <https://doi.org/10.1080/00140139.2020.1774080>, 63(10), 1207-1220. <https://doi.org/10.1080/00140139.2020.1774080>
- Sauro, J. (2015). SUPR-Q: A comprehensive measure of the quality of the website user experience. *Journal of usability studies*, 10(2).
- Sauro, J., & Zarolia, P. (2017). SUPR-Qm: a questionnaire to measure the mobile app user experience. *Journal of Usability Studies*, 13(1), 17-37.
- Savela, N., Turja, T., & Oksanen, A. (2018). Social Acceptance of Robots in Different Occupational Fields: A Systematic Literature Review. *International Journal of Social Robotics*, 10(4), 493-502. <https://doi.org/10.1007/S12369-017-0452-5/TABLES/3>
- Schaefer, K. E. (2016). Measuring trust in human robot interactions: Development of the “trust perception scale-HRI”. En *Robust intelligence and trust in autonomous systems* (pp. 191-218). Springer.

- Schaefer, K. E., Straub, E. R., Chen, J. Y. C., Putney, J., & Evans, A. W. (2017). Communicating intent to develop shared situation awareness and engender trust in human-agent teams. *Cognitive Systems Research*, 46, 26-39. <https://doi.org/10.1016/J.COGLYS.2017.02.002>
- Schillaci, G., Bodiřoža, S., & Hafner, V. V. (2013). Evaluating the Effect of Saliency Detection and Attention Manipulation in Human-Robot Interaction. *International Journal of Social Robotics 2013 5:1*, 5(1), 139-152. <https://doi.org/10.1007/S12369-012-0174-7>
- Schmidtler, J., Bengler, K., Dimeas, F., & Campeau-Lecours, A. (2017). A questionnaire for the evaluation of physical assistive devices (quead): Testing usability and acceptance in physical human-robot interaction. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 876-881.
- Schmidtler, J., Knott, V., Hölzel, C., & Bengler, K. (2015). Human Centered Assistance Applications for the working environment of the future. *Occupational Ergonomics*, 12(3), 83-95. <https://doi.org/10.3233/OER-150226>
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4 (6), 103-108.
- Schrepp, M., & Thomaschewski, J. (2019). Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia & Artificial Intelligence*, 5(7). <https://doi.org/10.9781/ijimai.2019.06.006>
- Schwabe, H., & Castellacci, F. (2020). Automation, workers' skills and job satisfaction. *PLoS ONE*, 15(11 November). <https://doi.org/10.1371/JOURNAL.PONE.0242929>
- Sebastián García, O., & del Hoyo Delgado, M. A. (2002). *La carga mental de trabajo*.
- Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable eda device. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 410-417. <https://doi.org/10.1109/TITB.2009.2036164>
- Shawn Lawton, H. (2003). *WAI Site Usability Testing Questions*. <https://www.w3.org/WAI/EO/Drafts/UCD/questions.html>
- Simões, A. C., Pinto, A., Santos, J., Pinheiro, S., & Romero, D. (2022). Designing human-robot collaboration (HRC) workspaces in industrial settings: A systematic literature review. *Journal of Manufacturing Systems*, 62, 28-43. <https://doi.org/10.1016/J.JMSY.2021.11.007>
- Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., & Friedman, D. (2006). Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 15(5), 553-569. <https://doi.org/https://doi.org/10.1162/pres.15.5.553>
- Slater, M., Usoh, M., & Steed, A. (1994). Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 3(2), 130-144. <https://doi.org/10.1162/PRES.1994.3.2.130>
- Stein, A., Yotam, Y., Puzis, R., Shani, G., & Taieb-Maimon, M. (2018). EEG-triggered dynamic difficulty adjustment for multiplayer games. *Entertainment Computing*, 25, 14-25. <https://doi.org/https://doi.org/10.1016/j.entcom.2017.11.003>
- Strube, M. J., & Werner, C. (1984). Personal space claims as a function of interpersonal threat: The mediating role of need for control. *Journal of Nonverbal Behavior*, 8, 195-209.

- Sun, S., Zheng, X., Gong, B., Paredes, J. G., & Ordieres-Meré, J. (2020). Healthy Operator 4.0: A Human Cyber–Physical System Architecture for Smart Workplaces. *Sensors 2020, Vol. 20, Page 2011, 20(7)*, 2011. <https://doi.org/10.3390/S20072011>
- Svarny, P., Tesar, M., Behrens, J. K., & Hoffmann, M. (2019). Safe physical HRI: Toward a unified treatment of speed and separation monitoring together with power and force limiting. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 7580-7587.
- Svendsen, G. B., Johnsen, J. A. K., Almås-Sørensen, L., & Vittersø, J. (2013). Personality and technology acceptance: the influence of personality factors on the core constructs of the Technology Acceptance Model. *Behaviour & Information Technology*, 32(4), 323-334. <https://doi.org/10.1080/0144929X.2011.553740>
- Taffese. (2017). *A Review of Using EEG and EMG Psychophysiological Measurements in User Experience Research*.
- Tang, G., & Webb, P. (2018). The Design and Evaluation of an Ergonomic Contactless Gesture Control System for Industrial Robots. *Journal of Robotics*, 2018. <https://doi.org/10.1155/2018/9791286>
- Teverovsky, D. (s. f.). *A Critical Review of Eye Tracking in UX*.
- Thakkar, D. (2018). *Biometric Devices: Cost, Types and Comparative Analysis*. <https://www.bayometric.com/biometric-devices-cost/>
- Thrun, S. (2004). Toward a framework for human-robot interaction. *Human–Computer Interaction*, 19(1-2), 9-24.
- Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, 42(4), 253-264. <https://doi.org/10.1080/00207590701396674>
- Tsai, C.-S., Hu, J.-S., & Tomizuka, M. (2014). Ensuring safety in human-robot coexistence environment. *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4191-4196. <https://doi.org/10.1109/IROS.2014.6943153>
- Tsarouchi, P., Makris, S., & Chryssolouris, G. (2016). Human–robot interaction review and challenges on task planning and programming. *International Journal of Computer Integrated Manufacturing*, 29(8), 916-931. <https://doi.org/https://doi.org/10.1080/0951192X.2015.1130251>
- Tsui, K. M., Abu-Zahra, K., Casipe, R., M’Sadoques, J., & Drury, J. L. (2010). Developing heuristics for assistive robotics. *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 193-194.
- Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, 9(5), 497-503. <https://doi.org/10.1162/105474600566989>
- Väänänen-Vainio-Mattila, K., Roto, V., & Hassenzahl, M. (2008). Towards practical user experience evaluation methods. *Meaningful measures: Valid useful user experience measurement (VUUM)*, 19-22.
- Veling, L., & McGinn, C. (2021). Qualitative Research in HRI: A Review and Taxonomy. *International Journal of Social Robotics*, 13(7), 1689-1709. <https://doi.org/10.1007/S12369-020-00723-Z/FIGURES/3>
- Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39(2), 273-315.

- Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55, 248-266. <https://doi.org/10.1016/J.MECHATRONICS.2018.02.009>
- Villani, V., Sabattini, L., Czerniak, J. N., Mertens, A., Vogel-Heuser, B., & Fantuzzi, C. (2017). Towards modern inclusive factories: A methodology for the development of smart adaptive human-machine interfaces. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, 1-7. <https://doi.org/10.1109/ETFA.2017.8247634>
- Virzi, R. A. (1992). Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(4), 457-468. <https://doi.org/10.1177/001872089203400407>
- von der Pütten, A. R., & Bock, N. (2018). Development and Validation of the Self-Efficacy in Human-Robot-Interaction Scale (SE-HRI). *ACM Transactions on Human-Robot Interaction (THRI)*, 7(3). <https://doi.org/10.1145/3139352>
- von Saucken, C., & Gomez, R. (2014, octubre). Unified User Experience Model enabling a more comprehensive understanding of emotional experience design. *ds.). proceedings of the Colors of Care: the 9th international Conference on design & emotion*.
- Wang, J., & Senecal, S. (2007). Measuring perceived website usability. *Journal of Internet Commerce*, 6(4), 97-112.
- Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X. V., Makris, S., & Chryssolouris, G. (2019). Symbiotic human-robot collaborative assembly. *CIRP Annals*, 68(2). <https://doi.org/10.1016/j.cirp.2019.05.002>
- Wang, X. V., Kemény, Z., Váncza, J., & Wang, L. (2017). Human–robot collaborative assembly in cyber-physical production: Classification framework and implementation. *CIRP Annals - Manufacturing Technology*, 66(1), 5-8. <https://doi.org/10.1016/j.cirp.2017.04.101>
- Wang, X. V., Seira, A., & Wang, L. (2018). Classification, personalised safety framework and strategy for human-robot collaboration. *Proceedings of International Conference on Computers & Industrial Engineering, CIE*.
- Warta, S. F., Kapalo, K. A., Best, A., & Fiore, S. M. (2016). *Similarity, Complementarity, and Agency in HRI: Theoretical Issues in Shifting the Perception of Robots from Tools to Teammates; Similarity, Complementarity, and Agency in HRI: Theoretical Issues in Shifting the Perception of Robots from Tools to Teammates*. <https://doi.org/10.1177/1541931213601287>
- Watson, D., & Clark, L. A. (1999). *The PANAS-X: Manual for the positive and negative affect schedule-expanded form*.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063.
- Weiss, A., Bernhaupt, R., Tscheligi, M., & Yoshida, E. (2009). Addressing user experience and societal impact in a user study with a humanoid robot. *AISB2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction (Edinburgh, 8-9 April 2009)*, SSAISB, 150-157.
- Weiss, A., Wurhofer, D., Bernhaupt, R., Altmaninger, M., & Tscheligi, M. (2010). A methodological adaptation for heuristic evaluation of HRI. *19th International Symposium in Robot and Human Interactive Communication*, 1-6.

- Welfare, K. S., Hallowell, M. R., Shah, J. A., & Riek, L. D. (2019). Consider the human work experience when integrating robotics in the workplace. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 75-84.
- Wibowo, R. M., Erna, P. A., & Hidayah, I. (2017). Heuristic evaluation and user testing with ISO 9126 in evaluating of decision support system for recommendation of outstanding marketing officer. *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*, 454-458.
- Wiltshire, T. (2013). *Towards Modeling Social-Cognitive Mechanisms in Robots to Facilitate Human-Robot Teaming Related papers*. <https://doi.org/10.1177/1541931213571283>
- Wiltshire, T. J., Warta, S. F., Barber, D., & Fiore, S. M. (2017). Enabling robotic social intelligence by engineering human social-cognitive mechanisms. *Cognitive Systems Research*, *43*, 190-207. <https://doi.org/10.1016/J.COVSYS.2016.09.005>
- Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and Industry 5.0—Inception, conception and perception. *Journal of Manufacturing Systems*, *61*, 530-535. <https://doi.org/10.1016/J.JMSY.2021.10.006>
- Ya-feng, N., Jin, L., Jia-qi, C., Wen-jun, Y., Hong-rui, Z., Jia-xin, H., Lang, X., Jia-hao, W., Guo-rui, M., Zi-jian, H., Cheng-qi, X., Xiao-zhou, Z., & Tao, J. (2022). Research on visual representation of icon colour in eye-controlled systems. *Advanced Engineering Informatics*, *52*, 101570. <https://doi.org/10.1016/J.AEI.2022.101570>
- Yan, S., Ding, G., Li, H., Sun, N., Guan, Z., Wu, Y., Zhang, L., & Huang, T. (2017). Exploring audience response in performing arts with a brain-adaptive digital performance system. *ACM Transactions on Interactive Intelligent Systems*, *7*(4), 1-28. <https://doi.org/https://doi.org/10.1145/3009974>
- Yanco, H. A., & Drury, J. (2004). Classifying human-robot interaction: An updated taxonomy. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, *3*. <https://doi.org/10.1109/ICSMC.2004.1400763>
- Yang, L., Zhang, H., Shen, H., Huang, X., Zhou, X., Rong, G., & Shao, D. (2021). Quality Assessment in Systematic Literature Reviews: A Software Engineering Perspective. *Information and Software Technology*, *130*, 106397. <https://doi.org/10.1016/J.INFSOF.2020.106397>
- Yang, T., Linder, J., & Bolchini, D. (2012). DEEP: design-oriented evaluation of perceived usability. *International Journal of Human-Computer Interaction*, *28*(5), 308-346. <https://doi.org/10.1080/10447318.2011.586320>
- Young, J. E., Sung, J., Volda, A., Sharlin, E., Igarashi, T., Christensen, H. I., & Grinter, R. E. (2011). Evaluating human-robot interaction: Focusing on the holistic interaction experience. *International Journal of Social Robotics*, *3*(1), 53-67. <https://doi.org/10.1007/S12369-010-0081-8>
- Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49-54.
- Zaki, T., Islam, M. N., & Zaki, T. (2021). Neurological and physiological measures to evaluate the usability and user-experience (UX) of information systems: A systematic literature review. *Computer Science Review*, *40*, 100375. <https://doi.org/10.1016/J.COSREV.2021.100375>
- Zickler, C., Halder, S., Kleih, S. C., Herbert, C., & Kübler, A. (2013). Brain painting: Usability testing according to the user-centered design in end users with severe motor paralysis. *Artificial Intelligence in Medicine*, *59*(2), 99-110. <https://doi.org/https://doi.org/10.1016/j.artmed.2013.08.003>

Złotowski, J., Yogeeswaran, K., & Bartneck, C. (2017). Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies*, 100, 48-54. <https://doi.org/10.1016/j.ijhcs.2016.12.008>

Anexo A

Questionario HUROY

Por favor, haga su evaluación ahora.

Valoramos sus comentarios sobre su reciente interacción con el robot. Para ayudarnos a comprender mejor su experiencia, le agradecemos que dedique unos minutos a rellenar el siguiente cuestionario. El cuestionario consiste en una serie de afirmaciones, y le pedimos que indique su nivel de acuerdo seleccionando un círculo que se ajuste a su opinión. Los círculos ofrecen una gama de gradaciones que van desde "Totalmente en desacuerdo" a "Totalmente de acuerdo", lo que le permite expresar sus pensamientos con mayores matices. Su opinión es crucial para ayudarnos a mejorar, así que le pedimos que sea sincero y reflexivo en sus respuestas.

Ejemplo:

1. Usar el robot es fácil.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Totalmente de acuerdo				

Esta respuesta significaría que no está de acuerdo con la afirmación anterior.

Déjate llevar por tus instintos y toma una decisión rápida. No lo pienses demasiado, confía en tu instinto y danos tu impresión genuina. Incluso si no estás del todo convencido o el robot no parece encajar perfectamente en una categoría determinada, simplemente marca un círculo en cada línea.

Tu opinión es lo más importante, y no hay respuestas correctas o incorrectas. Así que adelante, ¡siéntete libre y deja que tus pensamientos fluyan!

Utilidad percibida

1. El robot mejora mi rendimiento en la tarea.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

2. El robot me ayuda a ser más productivo.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

3. El robot me ayuda a completar la tarea en menos tiempo.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

4. El robot me da más control sobre las tareas.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

5. El robot me ayuda a realizar las tareas más fácilmente.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

6. El robot satisface mis necesidades.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

Facilidad de uso percibida

1. Usar el robot es fácil.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

2. Usar el robot simple.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

3. El robot es amigable.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

4. Usar el robot me ayuda.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

5. Puedo usar el robot siempre con éxito.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

6. Puedo utilizar el robot sin instrucciones escritas.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

7. El modo de control del robot requiere el menor número de pasos posible para lograr lo que quiero hacer con él.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

8. El modo de control del robot me permite recuperarme de los errores rápida y fácilmente.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

9. Conseguir un objetivo concreto con la ayuda del robot no será un problema para mí.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

Seguridad percibida

1. Me siento seguro/ra cuando utilizo el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

2. Durante la tarea, confiaba en que el robot no me haría daño.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

3. *Durante la tarea, me sentía ansioso/sa mientras utilizaba el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

4. *Durante la tarea, experimenté incomodidad al utilizar el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

Controlabilidad

1. El robot es fácil de controlar.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

2. Durante la tarea, he sentido que tenía el control del robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

3. Tengo mucha confianza en mi capacidad para controlar el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

4. Puedo configurar el robot según mis necesidades.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

5. Puedo ajustar el robot para que me ayude a realizar mis tareas.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

6. Puedo resolver problemas con la ayuda del robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

7. Si el robot está haciendo algo mal, podría encontrar una manera de cambiar su comportamiento.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

8. Podría desplegar el robot de una manera específica para ahorrar tiempo.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

Aprendibilidad

1. Puedo aprender fácilmente a utilizar el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

2. Puedo recordar fácilmente como se utiliza el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

3. He adquirido destreza con el robot fácilmente.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

Actitud e intención de uso

1. Podría utilizar el robot en mi comportamiento.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

2. Podría conseguir que un robot realizara una tarea específica.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

3. Si tengo la opción de elegir, prefiero trabajar con el robot.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

4. Creo que utilizaría el robot en futuras tareas.

	1	2	3	4	5	6	7	
Totalmente desacuerdo	<input type="radio"/>	Totalmente de acuerdo						

5. Durante la tarea, confiaba en que el robot hiciera lo correcto en el momento adecuado.

	1	2	3	4	5	6	7	
	<input type="radio"/>							

Totalmente desacuerdo Totalmente de acuerdo

Satisfacción

1. Estoy satisfecho/cha con el robot.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

2. Durante la tarea, el robot y yo trabajamos bien juntos.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

3. El robot funciona como yo quiero que funcione.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

4. El uso del robot es armonioso.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

5. * El robot me hace sentir incómodo/da.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

6. Me gusta usar el robot.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

7. Me siento cómodo/cómoda con el robot.

1 2 3 4 5 6 7
Totalmente desacuerdo Totalmente de acuerdo

Anexo B

Listado de heurísticos HEUROBOX

HEUROBOX

Heuristic Robots Experience

What is HEUROBOX?

HEUROBOX is a heuristic evaluation to assess Human-Robot Interaction (HRI) in industrial environments. It is a method of usability evaluation based on the inspection of different sections. In short, it consists of checking the quality of a set of principles called heuristic principles, i.e., checklists items. It is an agile and low resources method, which is known to detect usability errors.

General structure

The structure of each section follows a classification system of heuristics. HEUROBOX contains four categories of heuristics: Safety, Ergonomics, Functionality and Interface. Each category also has sub-categories.

In addition, there are two types of evaluation:

- The **Basic evaluation** is a simpler and quicker evaluation consisting of a total of 84 heuristics.
- The **Advanced evaluation** is a more complex evaluation but allows more usability errors to be found. It consists of a total of 228 heuristics.

Thus, there are general principles related to each section, items that correspond to the basic level and items that correspond to the advanced level. Each of the categories is made up of sub-categories, each of which has a basic evaluation section and an advanced evaluation section. If we want to carry out the advanced evaluation, both the basic and advanced evaluation heuristics must be filled in.

Heading to locate the user - **Category and subcategory**

Subcategory and brief description of what is to be assessed

Heuristics **Basic evaluation**



Heuristics **Advanced evaluation**

1

Safety.

Safety is the main challenge to be addressed in HRI. In this context safety and perceived risk are a determining factor, as they have a direct impact on the performance and the emotions of the person.

2

Ergonomics.

New robotic environments must be able to prevent worker discomfort as well as aggravation of diseases related to postural or physical conditions in workspaces, as musculoskeletal disorders constitute the largest category of work-related diseases in many industrial countries. Cognitive ergonomics is the field of study that focuses on how well the use of a product matches the cognitive capabilities of users.

3

Functionality.

It refers to the functions the robot has based on efficiency of the task and how well it is implemented, covers usability in terms of functions available, how system information is presented to the user, error handling methodologies and assistance.

4

Interfaces.

Collaborative robots can be multimodal, i.e., they usually consist of more than one type of interface. Interfaces can generate different types of communication, from graphical language, or voice-based communication to gesture-based dialogue. In addition, the types of interfaces change with the above in mind. For example, graphical communication can take place using specific devices, while voice-based communication can use natural language interfaces and gesture-based communication can use cameras suitable for tracking human hands.

How does it work?

- 1. Place the evaluator in context:** The evaluator should inquire about the profile of the user and the function that the robot has, as this is the basis for a good evaluation.
- 2. Fill in the Technical Datasheet:** Before starting the evaluation, it is advisable to fill in the technical datasheet. This section contains technical data of the system. A series of questions are also asked to determine how the person-robot relationship occurs (coexistence, cooperation, collaboration...). Finally, a series of tasks are proposed, which it is advisable for the evaluator to carry out before starting the evaluation, as this will give him/her a better understanding of the system to be evaluated.
- 3. Execution of the evaluation:** Once you have chosen the level of evaluation, proceed with it. The template must be filled in category by category. It should be noted that for each item there are three boxes: "yes", "no" and "not applicable" (in case the item is not applicable or does not correspond to the robotic system in question). The items have been proposed in such a way that the appropriate response is positive. So, after all the checks have been carried out, the answers marked "yes" will correspond to the number of items that the robotic system adequately fulfils, while the "no" answers will correspond to aspects that need to be improved. Finally, there is an area for scoring and comments. In the **interfaces category**, the sub-categories correspond to the corresponding interface types. Therefore, only the heuristics of the interfaces of the system to be evaluated have to be filled in.
- 4. Analysis of the results.** At the end of each category, there is a section to count the results. The "yes" for each of the categories must be counted over the number of items that have been evaluated. The "Na" items do not have to be counted as they are not applicable to the system in question. The heuristics that have been marked as "No" are the aspects to be improved.

Technical datasheet.

Before starting the evaluation, fill in this section.

Date:

Robot brand:

Robot model:

Year of robot:

	Yes	No
The person and the robot have a shared workspace .	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have direct contact .	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have the same task .	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have simultaneous process .	<input type="checkbox"/>	<input type="checkbox"/>
The person and the robot have sequential process .	<input type="checkbox"/>	<input type="checkbox"/>

Choose the **types of interface** the robot has:

Visual Voice Haptic Gesture

Choose the **context** the robot has:

Laboratory Real

Before starting the assessment, do the following **tasks**:

Make an observation of the workspace.

Execute all robot functionalities.

Safety.

In this section you will find general assessment in the Basic evaluation and specific sub-categories in the Advanced evaluation. This are all the sub-categories in this section:

- 1) General
- 2) Motion planning
- 3) Robot systems
- 4) Organizational measures

01

01.1	General	Page 8
01.2	Motion planning	Page 9
01.3	Robot systems	Page 10
01.4	Organizational measures	Page 11

GENERAL

This section assesses generic aspects of system security.

Basic evaluation

	Yes	No	Na
1. The system minimizes all the general hazards related to human-assembly parts interaction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system minimizes specific mechanical hazards related to the entrapment of human body parts.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system minimizes specific mechanical hazards related to human – assembly parts interaction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system minimizes specific mechanical hazards related to robot system parts falling.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system does not have a physical form that can induce injury.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system does not have behaviors that can induce injury.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The system provides fail-safe mechanism.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

MOTION PLANNING

This section assesses security aspects regarding Motion planning settings.

Advanced evaluation

	Yes	No	Na
1. The system allows to set trajectories in such a way human body parts will not be easily hit by the robot systems.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system allows to set trajectories in such a way human body parts will not be easily trapped between the robot systems and the elements of the workstation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system allows to limit velocities of moving parts.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system allows to limit momentum, mechanical power or energy as a function of masses and velocities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system has a safe virtual-plane-systems or space limiting functions which limit the robot to work in a defined volume.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system has speed limits for quasi-static contact.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The forces and torques of the system are limited via SW.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The system uses safety-rated soft axis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The systems speeds are limited for transient contact.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The robot must identify landmarks in the route description.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ROBOT SYSTEMS

Analyses the characteristics of the robot system and workspaces.

Advanced evaluation

	Yes	No	Na
1. The system has smooth surfaces.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system allows to manage energy absorption, enlarge energy transfer time and reduce impact forces (provide padding, cushioning or deformable components).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system allows to limit momentum, mechanical power or energy as a function of masses and velocities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system uses sensing to anticipate or detect contact.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The end-effector is designed to provide protection from hazards associated with the workpiece.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system avoids entrapment due to the moving cables of the robot systems.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The system avoids entrapment or jamming due to exposed parts of the robot systems.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The objects and obstacles into the system are highlighted.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ORGANIZATIONAL MEASURES

Analyses the organizational measures to ensure safe conditions.

Advanced evaluation

	Yes	No	Na
1. Robot system motion is signaled /highlighted.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The transition between collaborative operations and other kind of operations is signaled.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The performance of robotic system enables it to be monitored.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system allows the establishment of access routes (e.g. the paths taken by operators moving materials to the collaborative workspace).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

01 SAFETY

01.1 General: ____ / ____

01.2 Motion planning: ____ / ____

01.3 Robot Systems: ____ / ____

01.4 Organizational measures: ____ / ____

TOTAL SCORE: ____ / ____

Comments:

Ergonomics.

In this section you will find heuristic evaluation in two sub-categories:

- 1) Physical Ergonomics
- 2) Cognitive Ergonomics

02

02.1 Physical Ergonomics

Page 14

02.2 Cognitive Ergonomics

Page 16

PHYSICAL ERGONOMICS

Physical Ergonomics concerns with human anatomical, anthropometric, physiological, and biomechanical characteristics as they relate to physical activity.

Basic evaluation

	Yes	No	Na
1. The system physics is designed with ergonomics and accessibility in mind.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system minimizes the bio-mechanical overload of upper limbs related to repetitive tasks.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system minimizes the bio-mechanical overload of whole body related to manual lifting/lowering of objects.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system minimizes the bio-mechanical overload of head/neck/trunk/upper or lower limbs related to static or awkward working postures.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
5. The system avoids tasks which require the use of upper limbs for long time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system avoids tasks which require the elbows position above the shoulder level for quite all the time .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The system avoids tasks which require the use of moderate and continuous force.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The system avoids tasks which require force peaks.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system avoids tasks which require the need of grasping using the fingers tips (all kinds) for quite all the time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The system avoids tasks which require high frequency and similar movements of upper limbs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The system avoids tasks which require to maintain the workstation elements far to the body.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
12. The system avoids tasks which require a vertical displacement outside the range between hips and shoulders.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The system avoids tasks which require frequent body movements.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. The system reduces the weight and/or support heavy equipment, devices and, in general, every workstation elements manipulated by the operators.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The system avoids tasks which require an asymmetric posture of booth neck and trunk.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. The system avoids tasks which require unsupported trunk backward inclination or harsh flexion.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. The system avoids tasks which require neck extension or harsh flexion during the assembly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. The system avoids tasks which require unsupported head backward inclination or harsh inclination.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. The system avoids tasks which require a convex spinal curvature (if sitting).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. The system avoids tasks which require awkward upper arm postures.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. The system avoids tasks which require raised shoulder.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. The system avoids tasks which require unsupported upper arm elevation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. The system avoids tasks which require extreme elbow flexion/extension AND extreme forearm rotation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. The system avoids tasks which require extreme wrist deviation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. The system avoids tasks which require extreme knee flexion.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. The system avoids tasks which require knee not flexed in standing postures.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. The system avoids tasks which require not-neutral ankle position.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. The system avoids tasks which require kneeling or crouching.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. The system avoids which require very high knee angle (if sitting).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

COGNITIVE ERGONOMICS

Cognitive Ergonomics focuses on how well the use of a product matches the cognitive capabilities of users. It draws on knowledge of human perception, mental processing, and memory.

Basic evaluation

	Yes	No	Na
1. The system makes components well identifiable and distinguishable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system makes components easy to orient, locate and fasten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system makes the work intuitive (supports the formation of a mental model, reduces the choice reaction time, facilitates the leaning transfer, promotes similarity)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system maximizes the user's trust.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The UI is designed with ergonomics and accessibility in mind, ensuring the cobot UI is comfortable to work with for the necessary duration.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system maximizes operator psychological wellbeing and satisfaction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The system avoids high speed motions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The robot is as low as possible	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system informs operators about the robot speed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
10. The system makes menu items that need attention visually salient, do not attrat attention unnecesarily.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The system provide fused sensor information to avoid making the user fuse the data mentally.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
12. The system supports the user in understanding the connection between user actions and system response, for instance by providing feedback and using appropriate terminology.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The system ensures the robot performs in accordance with polite social etiquette.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. The system provides fused sensor information to lower the cognitive load on user.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The system reduces the number of assembly sub-systems.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. The system allows an immersive operator experience: sense of "being there".	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. The system implements smooth trajectories (which can be assimilate to natural human-arm motions).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. The system implements swing trajectories (not continuously straight).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. The system involves operators into the definition of layout and work activities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. The system avoids misalignment in operator and robot use of production resources (avoids inefficiency).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

02 ERGONOMICS

02.2 Physical Ergonomics: ____ / ____

02.2 Cognitive Ergonomics: ____ / ____

TOTAL SCORE: ____ / ____

Comments:

Functionality.

In this section you will find heuristics in five sub-categories:

- 1) System
- 2) Information
- 3) Task
- 4) Error handling
- 5) Assistance

03

03.1	System	Page 20
03.2	Information	Page 22
03.3	Task	Page 24
03.4	Error handling	Page 25
03.5	Assistance	Page 26

SYSTEM

Covers the usability of the system in terms of functions available.

Basic evaluation

	Yes	No	Na
1. The interaction with the robot is consistent and predictable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system supports the user in understanding the way it works.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system programs, trajectories and configurations are editable by the user.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system enables users to change the UI according to their needs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system provides fail-safe mechanisms.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system allows manual control to move around (mobile robotics).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
7. The user can decide the interaction level, settings and actions, i.e. to integrate human input or to make the program fully automatic (flexibility of interaction).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The system's software integration is reconfigured easily after hardware exchange; i.e. scalability of the architecture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system manages autonomously user attention across different modalities.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The system allows users to switch between simple and more complex ways of programming the cobot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The system manipulates relationship between robot and world.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The system provides assistance with autonomous modes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The interface switches modes automatically when necessary.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
14. Sensor and actuator capabilities are adequate for the system's expected tasks and environment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The system accommodates the ability to choose among access devices.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. The system provides multiple ways to access a function.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. The system is rigidly designed with the appropriate amount of interaction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. The system shows new functionalities as robotic systems gain new hardware and software components.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

INFORMATION

Defines the assessment of how system information is presented to users.

Basic evaluation

	Yes	No	Na
1. User can operate the robot using different channels for input and output.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. User can check and correct interaction information before submitting.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system informs the user regarding the robot's environment (spacial information) and configuration (indicators of robot health/state).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system allows users to access information required for the task, i.e. gives mission feedback.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system communicates to the user which task is being executed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The UI uses simple language to reduce mental processing.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The system avoids having the user to make mental translations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The system provides effective (precise and accurate) communication architecture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system provides with large video windows that assist in the success of the task.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The system use a single monitor for the interface.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The system gives continuous understanding of the overall mission and the moment-by-moment at a reasonable time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The system informs about progress towards completing the mission.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The system provides consistency of information; i.e. a match between system and real world information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
14. User can at all times query what the robot is doing or processing.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The system minimises the use of multiple windows; i.e. supports multiple information in a single window, if possible (e.g., multi robot systems).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. The interface shows consistency between robot behavior and what the operator has been led to believe based on the interface (actions).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. When multiple robots are available, the system allows to use one to view another.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. The system implicitly switches interfaces modality and autonomy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. The system conveys the information of the video stream with respect to robot orientation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. The system uses the principle of recognition over recall, improving users' situational awareness via attention management; i.e. helps directing the operator's focus of attention.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. The system provides option awareness to enable decision makers to know what courses of action are available, what their likelihoods of success are, and what their relative costs are.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. The system provides users with sufficient historical information to understand trends and make predictions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. The systems exploits existing systems' long-term and working memory as user support.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. The system provides knowledge in the interface so that people do not have to remember it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. The system allows retention of user's preferences.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. The system provides feedback and interaction according to user's technical abilities (skills)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. The systems shows awareness for the robot's autonomy-influenced behaviours.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
28. The system provides templates, instructions or other clues that indicate how the cobot can be interacted with.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

TASK

Based on the efficiency of the task and how well it is implemented.

Basic evaluation

	Yes	No	Na
1. Time does not interfere on users' ability to finish any interactive task with the robot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system helps minimising process length.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The sequence of tasks is efficient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The object recognition activities are simplified.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The feeding activities are simplified.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Handling activities are simplified.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Assembly activities are simplified.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
8. The system avoids unnecessary work on the user's side by minimising the number of steps required to achieve goals (e.g. recognition, feeding, handling or assembly), providing appropriate shortcuts.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The system enables reuse of their previous work or the work of others.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The system architecture allows adaptation of the system to different types of tasks and application scenarios.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The system exploits previous world knowledge if reasonable during task execution.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The system integrates robots early in team formation so that roles can be discovered and trust established quickly by human members.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ERROR HANDLING

Assess the actions to capture, discover and recover from an error in the system.

Basic evaluation

	Yes	No	Na
1. The system uses visual mechanisms to show errors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system logs or stores captured errors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The interface helps to recognise and diagnose user or robot errors, using appropriate descriptions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The system helps the user to recover from an error situation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system anticipates errors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Error messages or failures are not displayed using coding, machine or programming languages.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Error messages are displayed using human languages that are easy to understand, understood by the user.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
8. For any blinking component on robot's interface (lights, display contents, etc.) the blinking stops after a certain period, or can be switched off by user.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The interface helps to recover from errors made by the operator or the robot with explanations and steps.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The system provides a forgiving interface, allowing for reversible actions on the part of the operator or the robot as much as possible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The systems has the ability to self-inspect the robot's body for damages or entangled obstacles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The system presents enough information about the task environment so that the user can determine if some aspect of the world has contributed to the error.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
13. The system considers multiple source of errors (communication errors, input errors, mode errors...).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Before performing the action, save option, edit option, delete option, search option or other option user will be reminded last action through message box.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ASSISTANCE

Defines documentation or help (from to the system) methodologies.

Basic evaluation

	Yes	No	Na
1. User can ask for robot help or support at any given time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The system provides clear information and feedback.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The system uses standard rules for decision support.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Advanced evaluation

	Yes	No	Na
4. The system provides clear contextual help and documentation of available functionalities and possible errors (context-sensitive help).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system supports trial-and-error behaviors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The system provides the operator assistance in determining the most appropriate level of robotic autonomy (modality) at any given time (automation problem).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

03 FUNCTIONALITY

03.1 System: ____ / ____

03.2 Information: ____ / ____

03.3 Task: ____ / ____

03.4 Error handling: ____ / ____

03.5 Assistance: ____ / ____

TOTAL SCORE: ____ / ____

Comments:

Interfaces.

In this section you will find heuristics in five sub-categories:

- 1) General
- 2) Visual
- 3) Voice
- 4) Haptic
- 5) Gesture

04

- 04.1 General
- 04.2 Visual
- 04.3 Voice
- 04.4 Gesture
- 04.5 Haptic

Page 29

Page 31

Page 33

Page 34

Page 34

GENERAL

Covers the usability of the system in terms of functions available.

Basic evaluation

	Yes	No	Na
1. The UI supports the user in maintaining appropriate awareness of the system's state.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The UI makes it easy for the user to follow task execution by indicating previous, current and next steps.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The UI limits the number of options that are presented.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The UI attracts attention with the proper mechanisms.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The system and the UI behave in a consistent way.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The UI is easy and intuitive to use.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The design is intended for user's specificity, not the developer.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The UI grounds the information displayed with the reality.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The UI allows flexible, effective and efficient interactions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The UI makes it obvious what actions are available at any given moment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The UI follows real-world conventions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. System malfunctions are expressed in plain language (no codes).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The UI presents content appropriately.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. The system avoids the overload of the user's short-term memory, i.e., limiting the number of significant items (e.g., digits, letters, gestures) that must be remembered.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The UI is configured in different languages.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. The UI enables easy recovery from errors and offer guidance to the user on how they can correct it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
17. The UI provides ability to control and interrupt: The system should allow the user to interrupt if routed to a path they do not wish to follow.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. The system allows to use assistive technology to interact with the robot, such as screen reader, braille keyboards, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. The UI indicates which actions are currently possible and which ones are not (affordances).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. The UI tolerates minor user errors, prevent critical system errors, support undo and redo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. The UI gives users clear explanations of functionality and errors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. The UI supports easy editing of robot programs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. The UI ensures the way information is presented via different modalities consistently.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. The UI and system blend together so the UI is an extension of the user, the system and by proxy, the world.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. The physical embodiment of the system is pleasing in its intended setting. The UI do not contain information that is irrelevant or rarely needed.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. User can easily perceive and access robot's interfaces (hardware and software) components.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. If the system will be used over a lengthy period of time, the UI supports the evolution of system capabilities, such as sensor and actuator capacity, behavior changes and physical alteration.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Human and robot have a system of check-backs to ensure closed-loop communication (confirmation of reception).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

VISUAL

Analyses the interface based on information shown on displays or screens.

Advanced evaluation

	Yes	No	Na
1. The UI shows information in a way that is easy for the user to follow task progression.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The UI avoids a complex design, and makes use of simple graphics and icons.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The UI minimizes the use of multiple windows.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The UI presents the necessary information according to task, useful and relevant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The language of the interaction between the user and the system is in terms of words, phrases and concepts familiar to the user.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The design of the UI's items are usable, accessible, and aesthetic.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The information is presented in a clear and structured way, and use color, contrast and salience appropriately.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Fused information from different sensors are provided correctly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The UI has an aesthetic and minimalist design (simple use).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The UI displays the robot's body.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The UI presents automatic presentation of contextually-appropriate information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The UI presents clear information in appropriate form.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The UI prioritizes placement of information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. The UI allows the user to manipulate the information displayed and to store information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The UI presents sensor information that is easily understood and in a useful form.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	Yes	No	Na
16. The system displays programming functions at different levels of detail.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Windows occlusion hinders operation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. The UI presents information from multiple sensors presented in an integrated fashion.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. The UI has an easy transition to more in-depth information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. The UI uses efficient interaction language, using natural human cues.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. The GUI items are designed with usability, accessibility, and aesthetics in mind. The information is presented in a clear and structured way, and use color, contrast, and salience appropriately.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. The UI color is not the only way to indicate hardware controls, keys and labels of the robot. This also applies to software widgets (buttons, labels, etc.) or for information displayed on the robot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

VOICE

Analyses the voice interface.

Advanced evaluation

	Yes	No	Na
1. The UI allows the fewer number of steps that user-system dialog requires.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The UI ensures adequate system feedback. The system should always keep the user informed about what is going on through appropriate feedback within a reasonable time, providing, if necessary, confirmation of actions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The UI ensures high accuracy to minimize input errors. Recognition is important since errors degrade usability and lead to user frustration.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The UI is consistent and standard: Users should be able to maintain their focus on one interface or the link to a second interface (e.g. screen display) should be clear and consistent in operation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The VUI should speak in a natural way and adopt human-to-human speech conventions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The UI uses efficient interaction language, using natural human cues.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. The language of the interaction between the user and the system is in terms of words, phrases and concepts familiar to the user.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The dialog must be structured according to the four phases Introduction, Giving Directions, Confirmation, and Conclusion.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. The robot must interpret 'here' and 'there' depending on when they occur in the route description.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The robot must identify movement verbs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The delimitation problem must be solved by modeling the distance ranges of various expressions such as 'here', 'near', and 'far' depending on the environment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

HAPTIC

Assess the system input and output through touch interactions.

Advanced evaluation

	Yes	No	Na
1. The texture of the tactile object is perceptible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The tactile object does not need too much force.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The tactile objects level of vibration is significant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The duration of a haptic/tactile object is sufficient for the user to perceive changes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. The acceleration, the orientation speed of a tactile/haptic object is perceived and visualised.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Spatial attributes such as shape, location, and size exist for visual/graphic forms of interaction.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

GESTURE

Assess the system input and output through gesture interactions.

Advanced evaluation

	Yes	No	Na
1. The number of options involved in the interaction are as minimum as possible, to reduce the user's response times and errors.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The UI provides explicit hints to possible actions, guidelines on how to perform them and direct feedback on user's performance, to enforce learnability.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. The robot must solve the coordination problem, by asking the human to specify the first direction with a pointing gesture.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The robot must interpret the basic directions 'left', 'right', 'straight', and 'back'.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

04 INTERFACES

04.1 General: ____ / ____

04.2 Visual: ____ / ____

04.3 Voice: ____ / ____

04.4 Haptic: ____ / ____

04.5 Gesture: ____ / ____

TOTAL SCORE: ____ / ____

Comments:

Total scores.

01 SAFETY: ___ / ___

02 ERGONOMICS: ___ / ___

03 FUNCTIONALITY: ___ / ___

04 INTERFACES: ___ / ___

TOTAL SCORE: ___ / ___

Comments:

