07-002

# DEVELOPMENT AND VALIDATION OF HUMAN-ROBOT EXPERIENCE (HUROX) QUESTIONNAIRE FOR INDUSTRIAL COLLABORATIVE CONTEXTS

Apraiz, Ainhoa [(1)]; Lasa, Ganix [(1)]; Mazmela, Maitane [(1)]; Escallada, Oscar [(1)]; González de Heredia, Arantxa [(1)]

[(1)] Mondragon Unibertsitatea

An increasing number of robots are being implemented in Industry 5.0, which aims to put the well-being of the operators at the centre. From a human-centred design perspective, it is crucial to assess the perception and acceptance of robots. Questionnaires are a commonly used instrument for user experience assessment, as they allow an efficient quantitative measurement from the user's perspective. In the absence of questionnaires that assess user experience in an industrial robotic environment, the Human-Robot Experience (HUROX) questionnaire has been developed. using an empirical approach for the item selection to ensure the practical relevance of the items. Through a psychometric analysis where 15 experts have evaluated each of the items according to importance, necessity, relevance and clarity, the questionnaire has been validated. Therefore, the final version of the questionnaire is composed of a total of 41 items in 9 constructs (perceived usefulness, perceived ease of use, perceived safety, comfort, learning, controllability, attitude, trust and satisfaction). Therefore, this paper presents the HUROX questionnaire that allows to measure the perception and acceptance of the users in a simple way, while covering a complete impression of the interaction.

*Keywords:* Human-Robot Interaction (HRI);  User eXperience (UX); human factors; assessment; questionnaire; Industry 5.0

# DESARROLLO Y VALIDACIÓN DEL CUESTIONARIO HUMAN-ROBOT EXPERIENCE (HUROX) PARA CONTEXTOS COLABORATIVOS INDUSTRIALES

Cada vez es más común la implementación de los robots en la Industria 5.0, cuyo objetivo es situar el bienestar de los operarios en el centro. Desde una perspectiva de diseño centrado en las personas, es fundamental evaluar la percepción y aceptación de los robots. Los cuestionarios son un instrumento de uso común para la evaluación de la experiencia del usuario y aceptación tecnológica, ya que permiten una medición cuantitativa eficiente desde la perspectiva del usuario. Ante la carencia de cuestionarios que lo evalúen un entorno robótico industrial, se ha desarrollado el cuestionario Human-Robot Experience (HUROX). Mediante un análisis psicométrico donde 15 expertos han evaluado cada uno de los ítems según la importancia, necesidad, relevancia y claridad, se ha validado el cuestionario. Por lo tanto, la versión final del cuestionario se compone de un total de 41 ítems en 9 constructos (utilidad percibida, facilidad de uso percibida, seguridad percibida, comodidad, aprendizaje, controlabilidad, actitud, confianza y satisfacción). Por lo tanto, este trabajo presenta el cuestionario HUROX que permite medir la percepción y la aceptación de las personas usuarias de una forma sencilla, al tiempo que abarca una impresión completa de la interacción.

*Palabras clave:* Interacción Persona-Robot (HRI); Experiencia de Usuario (UX); factores humanos; evaluación; cuestionario; Industria 5.0

Agradecimientos: Los autores queremos agradecer el apoyo del Centro de Innovación de Diseño (DBZ) de Mondragon Unibertsitatea y a la Escuela Politécnica Superior de Mondragón.

# 1. Introduction

Industry 5.0 represents an industry's evolution towards greater integration of advanced technologies, especially in process automation and robotics. However, unlike its predecessors, Industry 5.0 places the well-being of humans as its central pillar (European Commission, 2021). Therefore, it is essential to ensure safe, efficient, and pleasant interaction between humans and robots. Two relevant frameworks for achieving this are Technology Acceptance and User Experience. By focusing on these aspects, the interaction between humans and robots can be made as beneficial as possible, thus improving human wellbeing in their working environment and in their daily jobs.

To start with, the Technology Acceptance Model (TAM) predicts the acceptance and use of information technologies based on perceived ease of use and perceived usefulness (Davis, 1989, 1993). The Human-Robot Collaboration Acceptance Model (HRCAM) proposed by Bröhl et al. (2019) found that job relevance was the most important predictor of perceived usefulness in the acceptance model for human-robot collaboration in an industrial context, followed by subjective norm, output quality, and result in demonstrability. The original TAM is transferrable to the domain of human-robot collaboration, with high correlation coefficients between perceived usefulness, perceived ease of use, behavioural intention, and use behaviour (Bröhl et al., 2019).

On the other hand, User Experience (UX) refers to a person's perceptions and responses resulting from the use or anticipated use of a product, system, or service, including emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours, and achievements that occur before, during, and after use (ISO 9241-210, 2019). UX research encompasses all factors that affect user interaction during the experience with a system or product (Hassenzahl, 2010; Wright et al., 2008). Effective human-robot interaction requires intuitive user interfaces (Dániel et al., 2014) that allow easy communication between humans and robots. To achieve this, it is necessary to define the intended interactions and purpose of information exchange between humans and robots based on both parties' application scope and functions (Driewer et al., 2007).

The success of HRI depends on the acceptance and satisfaction of its human users. The TAM and UX are two important frameworks that can aid in understanding and improving HRI in industrial contexts. By considering factors that influence user acceptance and experience, researchers and practitioners can develop more effective HRI systems that are better suited to the needs and preferences of human users.

In this sense, questionnaires have been widely used for Technology Acceptance and UX assessment, quantitatively providing valuable information on user perception. There are some case studies where UX is evaluated in industrial environments through questionnaires, such as the study conducted by Aranburu et al. (2020) or the study undertaken by Mazmela et al. (2019), which used the PANAS-X (Watson & Clark, 1999) and USE (Lund, 2001) questionnaires, respectively, after conducting a usability test. In both procedures, participants carried out usability tests with the HMI of industrial software, and after the test, they completed the indicated questionnaire.

In the context of human-robot interaction, von der Pütten & Bock (2018) developed and validated a new measure of self-efficacy. After conducting several experimental studies, they proposed a questionnaire consisting of 18 items. Participants had to rate the items on a six-point Likert scale (Likert, 1932). The survey by Apraiz et al. (2022) studied human factors resulting from Collaboration with an industrial robot in a virtual environment. To do so, they used the SE-HRI questionnaire (von der Pütten & Bock,

2018) and the questionnaire on UX in Virtual Reality proposed by Tcha-Tokey et al. (2016).

In the studies by Harriott et al. (2013), Aromaa et al. (2018) and Pantano et al. (2020), they used the NASA-TLX questionnaire (Hart & Staveland, 1988) to evaluate operator's workload in a human-robot context. On the other hand, in the studies by Schillaci et al. (2013) and Joosse et al. (2021), they used the Godspeed questionnaire (Bartneck et al., 2008) as it measures anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety.

In the study by Danielsson et al. (2017), they used the system usability scale (SUS) questionnaire (Brooke, 1996). The study by Almeida et al. (2020) used a questionnaire based on IBM Computer Usability Satisfaction Questionnaire (Lewis, 2009). Another self-created questionnaire was used in the study by Lasota & Shah (2015) to measure satisfaction with robots as teammates and determine perceived safety and comfort. In the study by Hietanen et al. (2020), they used a questionnaire to measure physical and mental stress; in the research by Beschi et al. (2020), they used a perceived risk questionnaire and a questionnaire on changes in planning configuration, in the study by Colim et al. (2021) they used a questionnaire on worker perception, in terms of perceived effort associated with the tasks and overall evaluation of the job.

There are still few applications in the field of human-robot interaction in industrial contexts, but they show that UX questionnaires present a great opportunity to evaluate. To a large extent because they collect constructs that can be interesting to apply in the context of Industry 5.0 and thus gain insight into user perception. By proposing agile tools, such as questionnaires, to evaluate the operator's interaction with machines, it is possible to identify the factors that must be maintained or improved to optimise this interaction and thus improve the processes that must occur in smart factories.

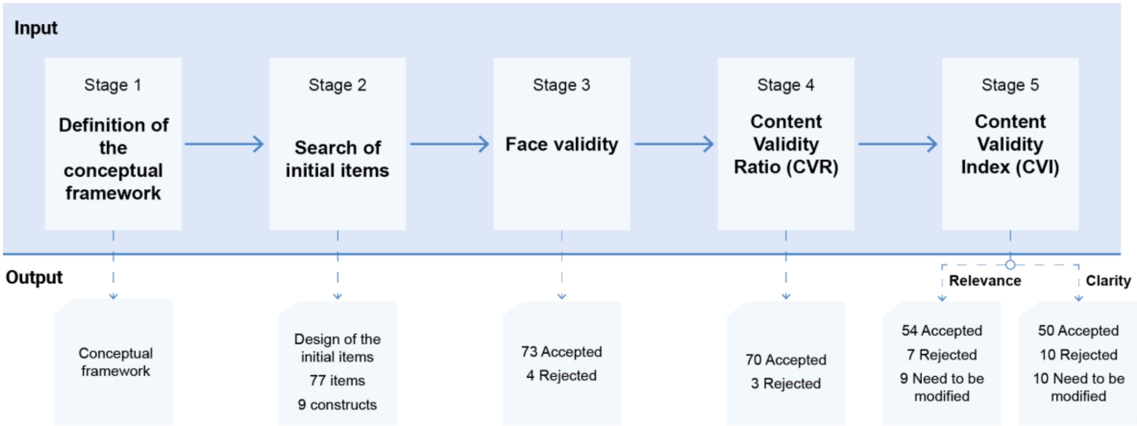
The integration of questionnaires with objective or qualitative tools can provide a more comprehensive understanding of both Technology Acceptance and UX. Although there is a rich literature on using questionnaires to evaluate Technology Acceptance and UX in Human-Computer Interaction (HCI) (Apraiz & Lasa, 2020), to the authors' knowledge, there is no questionnaire to assess UX and technology acceptance in an industrial collaborative context. Thus, creating an end-user questionnaire for assessing UX in human-robot interaction is a crucial step toward improving the design and acceptance of HRI systems in industrial contexts. By obtaining systematic and structured feedback from end-users, researchers and practitioners can identify issues and improve the overall UX, leading to more effective, efficient, and satisfying HRI systems. Additionally, an end-user questionnaire can help in comprehending end-users' attitudes toward the technology and identifying potential areas for improvement in the HRI system.

This paper presents the development and psychometric validation of the Human-Robot Experience (HUROX) questionnaire, which is an end-user questionnaire that assesses technology acceptance and UX through 41 items in 9 constructs (perceived usefulness, perceived ease of use, perceived safety, comfort, learnability, controllability, attitude, trust, and satisfaction).

## 2. Materials and methods

In the present study, questionnaire development and psychometric evaluation were performed in four stages according to the method proposed by Jahangiri et al. (2021). The following steps are explained (Figure 1):

**Figure 1: The followed methodology to design the HUROX questionnaire.**




### 2.1 Stage 1: Definition of the conceptual framework for the HUROX questionnaire

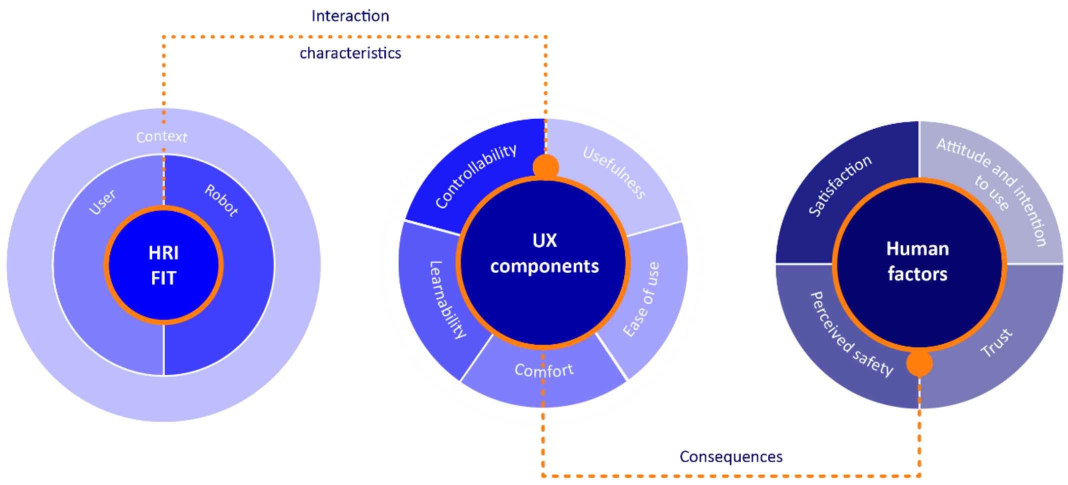
In the first stage, the conceptual framework was established. Building upon the CUE model proposed by Thüring & Mahlke (2007), which integrates experiential and instrumental aspects, the present study proposes the following conceptual framework (Figure 2). According to Thüring & Mahlke (2007), this framework addresses the characteristics of the system (in this case, the robot), the person, and the context. Therefore, the conceptual framework of this study argues that the interaction between a robot and a person is determined by a combination of the robot's characteristics, the user's characteristics, and contextual factors:

- Individual characteristics of the person, such as sociodemographic factors and intrinsic motivation, influence their expectations, skills, social norms, and preferences regarding robots.
- The robot's characteristics, including safety, task and robot programming, appearance, communication channels, and adjustability, are crucial in shaping the overall user experience.
- The context in which the interaction takes place, including the organizational and social context, also affects the usability and acceptability of robots.

In this sense, the "HRI FIT" between the three interaction characteristics has an impact on the UX components. For instance, the concept of usefulness denotes the extent to which an individual believes that using the robot would enhance their job performance (Davis, 1989). On the other hand, ease of use relates to the degree to which a person perceives using a particular system would be free of effort (Davis, 1989). With regard to comfort or ergonomics, we can divide between physical ergonomics, which concerns human anatomical, anthropometric, physiological, and biomechanical characteristics as they relate to physical activity (International Ergonomics Association, 2019); and cognitive ergonomics, which covers how well the use matches users' cognitive capabilities, including human perception, mental processing, and memory. Learnability signifies the extent to which users can successfully perform a task when they encounter an interface for the first time and the degree to which they can become proficient at that task with increasing repetitions of use (Joyce, 2019). Lastly, controllability refers to the degree to which a person feels in control of a technology and its actions, and the ability to modify its behaviour according to their needs or preferences (Sarker & Wells, 2003)

Therefore, if the fit between the interaction characteristics is good, the components of UX can lead to the improved overall experience and, which traduces in consequences on Human Factors such as attitude and intention to use, trust, perceived safety, and satisfaction, and ultimately facilitate the successful adoption and integration of robots in HRC settings.

**Figure 2: Conceptual framework for HUROX questionnaire.**




## 2.2 Stage 2: Design of the initial items

In the second stage, based on the literature, the items were designed for the constructs of the conceptual framework. To start the design of the initial items, a sizable collection of potential items was gathered for assessment. This collection was sourced from established questionnaires (Bartneck et al., 2008; Brooke, 1996; Lund, 2001; Schmidtler et al., 2017; von der Pütten & Bock, 2018), previous internal studies (Apraiz et al., 2022; Apraiz & Lasa, 2020; Aranburu et al., 2020; Mazmela et al., 2019, 2020), and brainstorming sessions. Subsequently, a process of curation was conducted to eliminate or rephrase items that were unsuitable for application in the domain of human-robot interaction.

### *2.2.1 Item sources*

After establishing the constructs that would form the questionnaire, the initial set of items was derived from a range of sources in the fields of psychology, human-computer interaction, human-robot interaction, user experience, and acceptance model. These are questionnaires that have been validated or used in experiments. The sources we used include the following questionnaires:

**Table 1: Item sources for the HUROX questionnaire.**

| Questionnaire | Reference | Validated | Constructs | Item construction | Nº of items |
|---|---|---|---|---|---|
| **SUS** | Brooke (1996) | Yes | Usability and learnability | Likert scale | 10 |
| **USE** | Lund (2001) | Yes | Learnability, usefulness, and ease of use | Likert scale | 30 |
| **Godspeed** | Bartneck et al. (2008) | Yes | Anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety | Semantic pair | 24 |
| **Self-generated** | Daniel et al. (2013) | No | Trust in automation | Likert scale | 19 |
| **Self-generated** | Lasota & Shah (2015) | No | Satisfaction with a robot as a teammate | Likert scale | 8 |
| **QUEAD** | Schmidtler et al. (2017) | Yes | Perceived usefulness, perceived ease of use, emotions, attitude, and comfort | Likert scale | 16 |
| **UEQ** | Hinderks et al. (2018) | Yes | Attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty | Semantic pair | 26 |

In addition, we have also relied on the following conceptual frameworks:

- The study by Berg et al. (2008) summarised key job findings on motivations and need, individual outcomes, positive experiences, unintended negative experiences, and resilience.
- The Human-Robot Collaboration Acceptance Model (HRCAM) developed by Bröhl et al. (2019), which is not a questionnaire per se, but a model of acceptance that defines key elements of human perception in HRC.

Table 2 displays the constructs that comprise the questionnaire, including their respective definitions and the total number of initial items for each construct that originated from the aforementioned questionnaires and conceptual frameworks.

**Table 2: Definition of the constructs that will compose the new questionnaire.**

| Construct | Definition | Nº of items |
|---|---|---|
| **Perceived usefulness** | The degree to which a person believes using a particular system would enhance their job performance (Davis, 1989). | 12 |
| **Perceived Ease of Use** | The degree to which a person believes using a particular system would be free of effort and easily learned (Davis, 1989). | 16 |
| **Perceived Safety** | The degree to which a particular system can acquire new knowledge or skills, as perceived by a person (Akalin et al., 2022). | 7 |
| **Ergonomics** | The degree to which a technology fits the user's physical and cognitive abilities and limitations, including aspects such as comfort, ease of use, and accessibility (Moroney & Lilienthal, 2008). | 3 |
| **Learnability** | The degree to which users can successfully perform a task when they encounter an interface for the first time, as well as the degree to which they can become proficient at that task with increasing repetitions of use (Joyce, 2019). | 5 |
| **Controllability** | The degree to which a person feels in control of a technology and its actions, and the ability to modify its behaviour according to their needs or preferences (Sarker & Wells, 2003). | 11 |
| **Attitude** | The user's overall positive or negative evaluation of a technology, affective and cognitive components. It is a predictor of the user's behavioural intention to use (Davis, 1989; Fishbein & Ajzen, 1977). | 7 |
| **Trust** | The degree to which a person believes that a technology will function as intended and will not harm them or others (Mayer et al., 1995). | 3 |
| **Satisfaction** | The degree to which a person is pleased with the use of a technology, as the result of the whole usefulness, ease of use and trust (Lee & Chung, 2009). | 13 |
| **TOTAL** | | **77** |

### *2.2.2 Item construction*

According to Lewis & Mayes (2014), there are two basic approaches to item construction for standardized questionnaires designed to assess UX. One is to construct agreement items (participants indicate the extent to which they agree with a given statement) such as QUEAD (Schmidtler et al., 2017), the USE (Lund, 2001), the SUS (Brooke, 1996); the other is to construct semantic pairs (participants position their ratings between opposing endpointments) such as the Godspeed questionnaire (Bartneck et al., 2008), the PANAS (Watson et al., 1988), AttrakDiff (Hassenzahl, 2001) or the UEQ (Hinderks et al., 2018).

In accordance with standard practice in survey research, the items were formulated as seven-point rating scales based on the Likert format as, according to Preston & Colman (2000), the most reliable scores were derived from this scale. The Likert scale is a widely adopted approach in social science research that permits the quantification of subjective attitudes or perceptions by asking respondents to indicate their level of agreement or disagreement with a statement using a scale that ranges from strongly disagree to agree strongly (Allen & Seaman, 2007). By utilizing the Likert scale format, the items were designed to elicit nuanced and quantifiable responses from participants, enabling a detailed analysis of their perceptions and attitudes toward the robotic system under investigation.

The initial version of the questionnaire consisted of 77 items organised into 9 constructs. The initial items are presented in Table 3.

### 2.3 Stage 3: Face validity

During the third stage, the questionnaire's validity was assessed subsequent to the initial formulation of the item content. The validity of an instrument is determined by its ability to accurately measure the construct it was designed to assess (Jahangiri et al., 2021). The face validity of each item was assessed using the Item Impact Score (IIS), whereby only items with an IIS score of 1,5 or higher were considered acceptable. A panel of 15 experts were invited to evaluate the face validity of the questionnaire by providing ratings of the importance of each item using a five-point Likert scale ranging from "not important" (1), "somewhat important" (2), "moderately important" (3), "important" (4), "very important" (5).

The IIS score for each item was calculated as the product of the percentage of experts who assigned a score of 4 or 5 to the item and the mean rating score of each item, according to the formula 1:

IIS = percentage of experts who scored a 4 or 5 x mean score for of each item. (1)

### 2.4 Stage 4: Content Validity Ratio (CVR)

Content Validity Ratio (CVR) depends on the logical analysis of a contest, and its determination is based on subjective and individual judgment. To measure this index, experts were asked to classify each item based on a range of three-point Likert scale (1=unnecessary, 2=useful, but not essential, 3=necessary). The CVR proposed by Lawshe (1975) is a linear transformation of a proportional level of agreement on how many "experts" within a panel rate an item "essential" or "useful" calculated in the following way (Ayre & Scally, 2014):

$$\boldsymbol{CVR} = \frac{ne-\frac{n}{2}}{\frac{n}{2}} \quad (2)$$

Where,

n= total number of experts.

ne= number of experts who chose the "necessary" or "useful, but not essential" option for the questionnaire item.

Based on the number of experts evaluating items, the minimum acceptable CVR value should be based on the values introduced by Lawshe (1975) for the appropriateness of content validity. In the present study 15 experts participated; therefore, if the CVR value of the item was equal to or greater than 0,49, the item was accepted.

### 2.5 Stage 5: Content Validity Index (CVI)

According to Yusoff (2019), establishing the content validity is vital to support the validity of an assessment tool such as questionnaires, especially for research purpose. To measure CVI, we first ask the experts to rate the items in terms of Relevancy from 1 (not relevant), 2 (relatively relevant), 3 (relevant), to 4(completely relevant). Then, experts had to rate the items in terms of clarity ranging from 1(not clear), 2 (relatively clear), 3 (clear), to 4 (very clear) were given. According to Lawshe (1975), the CVI is an item statistic useful in rejecting or retaining specific items. After items have been identified for inclusion in the final form, the content validity index (CVI) is computed for the whole test. The CVI is simply the mean of the CVR values of the retained items.

For each item, the I-CVI can be calculated by counting the number of experts who rated the item as 3 or 4 and dividing that number by the total number of experts, that is the proportion of agreement about the content validity of an item.

$$I - CVI = \frac{ne}{n} \quad (3)$$

Where,

n= total number of experts.

ne= number of experts who chose the “3” or “4” option for the questionnaire item.

## 3. Results

This section presents the design and psychometric evaluation results of the HUROX questionnaire are presented. First, the results of the face validity assessment of the questionnaire are presented. Content validity

### 3.1 Characterisation of experts

The assessment was made by 15 experts who demonstrated diverse professional backgrounds and areas of expertise (Table 3). The 60% the experts were men, and experts professional experience spanned from 2 to 24 years. The participants’ expertise encompassed a wide range of domains, including UX, Human-Robot Interaction, Human-Centred Design, Ergonomics, Digital Inclusion, Interaction Design, and Manufacturing. This breadth of expertise enabled the participants to provide valuable insights into the questionnaire’s development and ensure its relevance to potential end-users across a range of industrial contexts.

**Table 3. Characterisation of the experts.**

| Experts id | Gender | Expertise | Working years |
|---|---|---|---|
| **1** | Man | User Experience | 4 |
| **2** | Woman | User Experience in Human-Robot Interaction | 4 |
| **3** | Man | User Experience in game-based systems | 2 |
| **4** | Man | Research scholar | 4 |
| **5** | Man | User Experience and interaction | 4 |
| **6** | Man | Computer science | 15 |
| **7** | Man | User Experience and Ergonomics | 7 |
| **8** | Man | User Experience and Ergonomics | 24 |
| **9** | Woman | Human-centred design | 3 |
| **10** | Woman | Digital inclusion | 3 |
| **11** | Woman | UX and Technology Acceptance in industrial environments | 7 |
| **12** | Woman | UX and Strategic Design | 4 |
| **13** | Man | Operations Excellence, Manufacturing, Human centred design | 8 |
| **14** | Woman | Human-centred design | 3 |
| **15** | Man | Interaction design | 12 |

### 3.2 Face validity of the Questionnaire

At the outset of the investigation, the questionnaire comprised 77 items. In order to appraise its face validity, the IIS was employed as the assessment tool. It is widely acknowledged that an IIS of no less than 1.5 is deemed acceptable for a research item to demonstrate face validity (Jahangiri et al., 2021). Therefore, only those items with an IIS above 1.5 were considered acceptable for inclusion in the study. The outcomes of the face validity evaluation, presented in Table 4, detail the IIS attributed to each item. As indicated, 3 items were rejected, PEU6, PEU11, and C8.

### 3.3 Content Validity Ratio of the Questionnaire

The results of CVR are shown in Table 4. According to Lawshe (1975), the acceptable CVR for 15 experts is 0,49. Accordingly, the items with a CVR of more than 0,49 are consistent and acceptable in terms of content with the objectives of the questionnaire. The results show that 4 items were rejected, concretely, PEU4, PEU7, PS3, and L5.

### 3.4 Content Validity Index (CVI) of the Questionnaire

The Content Validity Index (CVI) was computed twice for the questionnaire, namely, (1) by asking the participants to rate the relevance of each item and (2) by seeking their feedback on the clarity of each of them.

#### *3.4.1 Relevance Item-level Content Validity Index (I-CVI) of the Questionnaire*

The results of the CVI assessment related to Relevance are presented in Table 4. Based on Jahangiri et al. (2021) recommendations, the minimum acceptable value of 0,79 was set for the I-CVI. Items with an I-CVI score between 0,7 and 0,79 were identified as requiring modification, while those with an I-CVI score of less than 0,7 were rejected. Based on the analysis, 7 items (PU8, PEU16, PS5, PS6, PS7, A4 an S4) were rejected, and 9 items (PU1, E2, L2, C9, A5, T2, S5, S10 and S13) were identified as requiring modification.

#### *3.4.2 Clarity Item-level Content Validity Index (I-CVI) of the Questionnaire*

The results of the CVI assessment related to Clarity are presented in Table 4. Based on Jahangiri et al. (2021) recommendations, as previously, the minimum acceptable value of 0,79 was set for the I-CVI. Items with an I-CVI score between 0,7 and 0,79 were identified as needing modification, while items with an I-CVI score less than 0,7 were rejected. The analysis resulted in the rejection of 10 items (PU1, PU9, PU11, PS5, C4, A4, A5, A6, T3, and S8) and the identification of 10 items requiring modification (PU2, PU4, E1, C6, C7, A3, T2, S4, S6, and S7).

## 4. Discussion and conclusions

The purpose of this study was to develop and validate a new end-user questionnaire that can be used to evaluate the human perception of UX and technology acceptance in human-robot interaction industrial contexts, the Human-Robot Experience (HUROX) questionnaire. Using the HUROX questionnaire can help designers and developers to identify areas where improvements can be made to enhance the UX and technology acceptance of the system. It can serve as a benchmarking tool for comparing the UX and technology acceptance of different human-robot interaction systems. This comparison can provide insights into which systems are more effective in satisfying the users' needs and preferences and identify areas where improvements can be made.

The psychometric analysis of the HUROX questionnaire with 41 items in 9 constructs indicated acceptable face validity and content validity. The items were distributed across various constructs, reflecting the multidimensional nature of the construct being measured. The final version of the questionnaire provides a comprehensive and reliable tool for assessing user perceptions of the target system. These items were distributed across various constructs, with perceived usefulness consisting of 7 items, perceived ease of use comprising 11 items, perceived safety including 3 items, comfort containing

1 item, learnability comprising 3 items, controllability consisting of 6 items, attitude including 3 items, trust comprising 1 item, and satisfaction consisting of 6 items.

However, in the psychometric evaluation of questionnaires, examining the reliability of the instrument is crucial. One commonly used method for measuring reliability is Cronbach's alpha coefficient (Cronbach, 1951). Therefore, as a future line of research, it is suggested that the reliability of the HUROX questionnaire be calculated by conducting a usability test in which participants interact with a human-robot system in an industrial environment and then complete the questionnaire. Additionally, it would be appropriate to analyse whether the 9 constructs that compose the questionnaire are balanced and accurately assessed. Another area for potential research is the cross-cultural validity of the questionnaire. Administering the questionnaire in different cultural and linguistic contexts could provide valuable insights into how user perceptions vary across different populations and regions, contributing to a more comprehensive understanding of human-robot interaction.

Despite the positive results, the sample size of the experts in this study is relatively small and may not be representative of the entire population of experts in the field. Furthermore, the study relied on self-reported data from the experts, which could be subject to bias or inaccuracies. Another limitation of the study is that it did not include experts from other related fields, such as psychology, anthropology, or sociology, which may limit the breadth of perspectives on UX.

In conclusion, the HUROX questionnaire is a valuable tool for assessing user perceptions of human-robot interaction systems in industrial settings. Nonetheless, future studies should aim to address the limitations and gaps identified in this research to strengthen the validity and generalizability of the questionnaire.

**Table 4: Results for the calculations of IIS, CVR and CVI of the questionnaire.**

| Construct | Code | Item | Stage 3 - Face validity | | Stage 4 - Content validity ratio | | Stage 5 - CVI | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Relevance | | Clarity | |
| | | | IIS | Result | CVR | Result | I-CVI | Result | I-CVI | Result |
| Perceived usefulness | PU1 | The robot is useful. | 3,36 | Accepted | 1,00 | Accepted | 0,73 | **Modify** | 0,67 | **Rejected** |
| | PU2 | The robot helps me be more effective. | 3,31 | Accepted | 0,87 | Accepted | 0,93 | Accepted | 0,73 | **Modify** |
| | PU3 | The robot helps me be more productive. | 3,81 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,87 | Accepted |
| | PU4 | The control mode of the robot enhances my working performance. | 3,58 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,73 | **Modify** |
| | PU5 | The robot enables me to accomplish the given task rapidly. | 2,67 | Accepted | 0,73 | Accepted | 0,80 | Accepted | 0,80 | Accepted |
| | PU6 | The robot saves me time when I use it. | 2,71 | Accepted | 1,00 | Accepted | 0,87 | Accepted | 0,87 | Accepted |
| | PU7 | The robot gives me more control over the activities in my daily job. | 2,62 | Accepted | 1,00 | Accepted | 0,87 | Accepted | 0,93 | Accepted |
| | PU8 | The robot makes the things I want to accomplish easier to get done. | 3,64 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,80 | Accepted |
| | PU9 | I was able to perform precise motions with the robot's control mode. | 2,32 | Accepted | 0,87 | Accepted | 0,67 | **Rejected** | 0,67 | **Rejected** |
| | PU10 | The robot meets my needs. | 3,36 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 0,87 | Accepted |
| | PU11 | The robot does everything I would expect it to do. | 3,41 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,67 | **Rejected** |
| | PU12 | Using the robot improves my performance at my job. | 4,36 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 0,93 | Accepted |
| Perceived ease of use | PEU1 | My interaction with the robot is easy. | 4,36 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,80 | Accepted |
| | PEU2 | The control mode of the robot is easy to use. | 4,67 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,80 | Accepted |
| | PEU3 | The control mode of the robot is simple to use. | 2,79 | Accepted | 0,73 | Accepted | 0,87 | Accepted | 0,80 | Accepted |
| | PEU4 | Using the robot is effortless. | 2,67 | Accepted | 0,47 | **Rejected** | - | - | - | - |
| | PEU5 | The control mode of the robot is user-friendly. | 3,70 | Accepted | 0,73 | Accepted | 0,80 | Accepted | 0,87 | Accepted |
| | PEU6 | *The control mode of the robot is rigid and inflexible. | 1,46 | **Rejected** | - | - | - | - | - | - |
| | PEU7 | *This control mode of the robot feels cumbersome. | 1,81 | Accepted | 0,33 | **Rejected** | - | - | - | - |
| | PEU8 | The control mode of the robot is helpful to me. | 3,87 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,87 | Accepted |
| | PEU9 | I can use the robot successfully every time. | 2,40 | Accepted | 0,73 | Accepted | 0,80 | Accepted | 0,87 | Accepted |
| | PEU10 | I can use the robot without written instructions. | 3,76 | Accepted | 1,00 | Accepted | 0,87 | Accepted | 1,00 | Accepted |
| | PEU11 | Both occasional and regular users would like the robot. | 1,46 | **Rejected** | - | - | - | - | - | - |
| | PEU12 | I don't notice any inconsistencies as I use the robot. | 2,12 | Accepted | 0,73 | Accepted | 0,80 | Accepted | 0,80 | Accepted |
| | PEU13 | The robot's control mode requires the fewest steps possible to accomplish what I want to do with it. | 3,64 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,93 | Accepted |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PEU14 | The robot's control mode enables to recover from mistakes quickly and easily. | 4,04 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,93 | Accepted |
| | PEU15 | To achieve a specific goal with the assistance of a robot will not be a problem for me. | 2,71 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 0,80 | Accepted |
| | PEU16 | I could do easy adjustments on a robot by myself. | 2,98 | Accepted | 0,87 | Accepted | 0,67 | **Rejected** | 0,80 | Accepted |
| **Perceived safety** | PS1 | I feel safe while using the robot. | 4,87 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 0,93 | Accepted |
| | PS2 | I trusted the robot would not harm me. | 4,17 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 1,00 | Accepted |
| | PS3 | I do not mind if the robot works with me at a shared workstation. | 1,81 | Accepted | 0,47 | **Rejected** | - | - | - | - |
| | PS4 | *I feel anxious while using the robot. | 2,93 | Accepted | 0,87 | Accepted | 0,80 | Accepted | 0,80 | Accepted |
| | PS5 | *I feel agitated while using the robot. | 2,53 | Accepted | 0,60 | Accepted | 0,60 | **Rejected** | 0,67 | **Rejected** |
| | PS6 | The robot moved too fast for my comfort. | 2,88 | Accepted | 0,73 | Accepted | 0,60 | **Rejected** | 0,93 | Accepted |
| | PS7 | The robot came too close to me for my comfort. | 2,88 | Accepted | 0,73 | Accepted | 0,60 | **Rejected** | 0,93 | Accepted |
| **Comfort (Ergonomics)** | E1 | I feel physically uncomfortable in using this control mode. | 3,70 | Accepted | 0,73 | Accepted | 0,87 | Accepted | 0,73 | **Modify** |
| | E2 | I feel tense in using this control mode. | 3,52 | Accepted | 0,73 | Accepted | 0,73 | **Modify** | 0,80 | Accepted |
| | E3 | I would feel comfortable while interacting with the robot. | 4,17 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,80 | Accepted |
| **Learnability** | L1 | I could easily learn how to use a robot. | 3,99 | Accepted | 0,87 | Accepted | 0,80 | Accepted | 0,87 | Accepted |
| | L2 | I learned to use it quickly. | 4,04 | Accepted | 0,87 | Accepted | 0,73 | **Modify** | 0,80 | Accepted |
| | L3 | I easily remember how to use it. | 4,04 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 1,00 | Accepted |
| | L4 | I quickly became skillful with it. | 4,36 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,87 | Accepted |
| | L5 | I can use the robot if someone shows me how to do it first. | 3,98 | Accepted | 0,47 | **Rejected** | - | - | - | - |
| **Controllability** | C1 | The robot is easy to control. | 4,42 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 0,93 | Accepted |
| | C2 | I had the overall control of the robot. | 4,42 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 1,00 | Accepted |
| | C3 | I am very confident in my abilities to control a robot. | 3,87 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 1,00 | Accepted |
| | C4 | I control the robot over job and meaning of work. | 2,24 | Accepted | 0,73 | Accepted | 0,80 | Accepted | 0,60 | **Rejected** |
| | C5 | I could set up a robot according to my wishes and my environment. | 3,09 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,80 | Accepted |
| | C6 | I think I could adjust a robot the way that it could help me in my daily life. | 3,70 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,73 | **Modify** |
| | C7 | If I should solve a problem with the assistance of a robot, I could do that. | 2,58 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 0,73 | **Modify** |
| | C8 | I could teach a robot something if I would try hard enough. | 1,49 | **Rejected** | - | - | - | - | - | - |
| | C9 | I could teach a robot to complete easy tasks. | 2,58 | Accepted | 1,00 | Accepted | 0,73 | **Modify** | 0,93 | Accepted |
| | C10 | If a robot is doing something wrong, I could find a way to change its behavior. | 4,04 | Accepted | 0,87 | Accepted | 1,00 | Accepted | 0,93 | Accepted |
| | C11 | I could deploy a robot in a specific way to save time. | 3,31 | Accepted | 0,73 | Accepted | 0,87 | Accepted | 0,80 | Accepted |
| **Attitude** | A1 | I could use a robot in daily life. | 3,47 | Accepted | 1,00 | Accepted | 1,00 | Accepted | 0,93 | Accepted |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A2 | I could get a robot to perform a specific task. | 3,87 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,93 | Accepted |
| | A3 | If I could choose whether the robot supports me at work, I would appreciate working with the robot. | 3,81 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,73 | **Modify** |
| | A4 | I prefer the robot to other machines in the industrial environment. | 2,93 | Accepted | 0,73 | Accepted | 0,67 | **Rejected** | 0,60 | **Rejected** |
| | A5 | The use of the robot is pertinent to my various job-related tasks. | 2,79 | Accepted | 1,00 | Accepted | 0,73 | **Modify** | 0,60 | **Rejected** |
| | A6 | I think that using this control mode is a good idea. | 1,88 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,67 | **Rejected** |
| | A7 | I think I would use this control mode in future task | 2,99 | Accepted | 0,73 | Accepted | 0,87 | Accepted | 0,80 | Accepted |
| **Trust** | T1 | I trusted the robot to do the right thing at the right time. | 4,04 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,87 | Accepted |
| | T2 | If I would use a robot, I would always know how and why it behaves like it does. | 3,57 | Accepted | 1,00 | Accepted | 0,73 | **Modify** | 0,73 | **Modify** |
| | T3 | I built a meaningful and helpful relationship with the robot. | 1,96 | Accepted | 0,87 | Accepted | 0,80 | Accepted | 0,47 | **Rejected** |
| **Satisfaction** | S1 | I am satisfied with the robot. | 4,23 | Accepted | 1,00 | Accepted | 0,93 | Accepted | 0,93 | Accepted |
| | S2 | The robot and I worked well together. | 3,87 | Accepted | 1,00 | Accepted | 0,87 | Accepted | 0,93 | Accepted |
| | S3 | It works the way I want it to work. | 4,36 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 0,87 | Accepted |
| | S4 | The robot aligns to my personal expectations. | 2,49 | Accepted | 1,00 | Accepted | 0,67 | **Rejected** | 0,73 | **Modify** |
| | S5 | I would recommend the use of the robot to a colleague. | 4,23 | Accepted | 0,87 | Accepted | 0,73 | **Modify** | 0,93 | Accepted |
| | S6 | The robot is pleasant to use. | 3,52 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,73 | **Modify** |
| | S7 | *The robot did not understand how I wanted to do the task. | 2,76 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 0,73 | **Modify** |
| | S8 | The robot kept getting in my way. | 3,31 | Accepted | 0,87 | Accepted | 0,80 | Accepted | 0,67 | **Rejected** |
| | S9 | Robots make me feel uncomfortable. | 3,36 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 0,87 | Accepted |
| | S10 | Robots make me additional stress. | 3,64 | Accepted | 0,87 | Accepted | 0,73 | **Modify** | 0,80 | Accepted |
| | S11 | I like using the robot. | 3,08 | Accepted | 1,00 | Accepted | 0,80 | Accepted | 0,93 | Accepted |
| | S12 | I feel comfortable the robot. | 3,64 | Accepted | 0,87 | Accepted | 0,87 | Accepted | 0,87 | Accepted |
| | S13 | *I feel anxious using the robot. | 3,36 | Accepted | 0,87 | Accepted | 0,73 | **Modify** | 0,93 | Accepted |

## 5. References

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

Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress*, *40*(7), 64–65.

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., 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*.

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

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

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

Berg, J. M., Dutton, J. E., & Wrzesniewski, A. (2008). What is job crafting and why does it matter. *Retrieved Form 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

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.

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

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Dániel, B., Korondi, P., Sziebig, G., & Thomessen, T. (2014). Evaluation of flexible graphical user interface for intuitive human robot interactions. *Acta Polytechnica Hungarica*, *11*(1), 135–151.

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

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.

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.

Driewer, F., Sauer, M., & Schilling, K. (2007). Discussion of Challenges for User Interfaces in Human-Robot Teams. *EMCR*.

European Commission, 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/

Fishbein, M., & Ajzen, I. (1977). *Belief, attitude, intention, and behavior: An introduction to theory and research*.

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

Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, *13*(4), 481–499.

Hassenzahl, M. (2010). Experience design: Technology for all the right reasons. *Synthesis Lectures on Human-Centered Informatics*, *3*(1), 1–95.

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

Hinderks, A., Schrepp, M., & Thomaschewski, J. (2018). *User Experience Questionnaire (UEQ)*. https://www.ueq-online.org/

International Ergonomics Association. (2019). *Definition and Domains of Ergonomics*. Ergonomics. https://doi.org/10.1080/00140139.2012.661087

ISO 9241-210. (2019). Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. In *International Standard: Vol. 2.Udgave*.

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

Joosse, 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

Joyce, A. (2019, October 20). *How to Measure Learnability of a User Interface*. https://www.nngroup.com/articles/measure-learnability/

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

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575.

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.

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.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*.

Lund, A. M. (2001). Measuring usability with the use questionnaire12. *Usability Interface*, *8*(2), 3–6.

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., 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, Nº. 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*.

Moroney, W. F., & Lilienthal, M. G. (2008). Human factors in simulation and training. *Human Factors in Simulation and Training. CRC Press*, 3–38.

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

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

Sarker, S., & Wells, J. D. (2003). Understanding mobile handheld device use and adoption. *Communications of the ACM*, *46*(12), 35–40.

Schillaci, G., Bodiroža, 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.

Tcha-Tokey, K., Loup-Escande, E., Christmann, O., & Richir, S. (2016). A questionnaire to measure the user experience in immersive virtual environments. *Proceedings of the 2016 Virtual Reality International Conference*, 1–5.

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

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

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.

Wright, P., Wallace, J., & McCarthy, J. (2008). Aesthetics and experience-centered design. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *15*(4), 1–21.

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49–54.