Usability Study through a Human-Robot Collaborative Workspace Experience

Alejandro Chacón [1,†] , Pere Ponsa [2,†] and Cecilio Angulo [3,4,*]

1 Department of Electrical and Electronics, Universidad de las Fuerzas Armadas 'ESPE', Quito 171103, Ecuador; lachacon@espe.edu.ec
2 Department of Automatic Control, Universitat Politècnica de Catalunya Barcelona Tech, 08019 Barcelona, Spain; pedro.ponsa@upc.edu
3 Intelligent Data Science and Artificial Intelligence Research Centre, 08034 Barcelona, Spain
4 Institut de Robòtica i Informàtica Industrial (CSIC-UPC), 08028 Barcelona, Spain
* Correspondence: cecilio.angulo@upc.edu
† These authors contributed equally to this work.

Abstract: The use of collaborative robots (cobots) in industrial and academic settings facilitates physical and cognitive interaction with operators. This framework is a challenge to determine how measures on concepts, such as usability, can be adapted to these new environments. Usability is a quality attribute prevalent in the field of human-computer interaction concerning the context of use and the measure of effectiveness, efficiency, and satisfaction of products and systems. In this work, the importance of the role of benchmarking usability with collaborative robots is discussed. The introduced approach is part of a general methodology for studying people and robots' performance in collaboration. It is being designed and developed on a concrete experience into a human-robot collaborative workspace. Outcomes from the study include a list of steps, resources, recommendations, and some customized questionnaires to obtain cobot-oriented usability analysis and case study results.

Keywords: usability; human-robot interaction; workload; benchmarking

1. Introduction

The Fourth Industrial Revolution is a new paradigm favoring the insertion of collaborative and autonomous robots in industrial environments. Insertion of new technologies means workplace redesign. Manual tasks carried out by human operators in industrial environments are being transformed into tasks now shared with collaborative robots, usually named cobots [1]. Cobots are complex machines working hand-in-hand with human operators. Their main task is to help and support human operators in the activities deemed necessary. These machines do not replace human labor, but they complement human capabilities [2].

Introduction of cobots is substantially modifying workplaces in industry. It would be recommended to analyze which usability engineering methods can be useful to measure this workplace modification, such as the ISO/TR 16982 standard [3]. Elements, such as ergonomics, safety, robot acceptance, human ability to understand the behavior of complex systems, trust, and tasks assignment between operator and robot, are modified [4,5]. Improvement in system performance should be demonstrated through human-robot frameworks design, new standard regulations, methods, metrics, and experimental tests [6–9].

Taking into account the project management cycle—requirement, specification, solution design, implementation, test, and validation—a human-centered design approach should allow introduce ergonomics guidelines and recommendations in early stages of the life cycle and usability methods throughout the entire product cycle management. In this moment, a question raises about the cost of adding usability into the project cycle

management. However, whether the decision has been made to transform the task and workspace with a collaborative robot, it is a key issue to focus on the human-centered approach. When adapting a small and medium enterprise for human-robot collaborative tasks, some usability methods should be taken into account: focus groups, thinking aloud, questionnaires, and expert evaluation. Human factor experts can choose and adapt the better methods oriented to a specific project. For instance, the thinking aloud method turns out to be a highly efficient method to capture aspects related to the cognitive/mental activities of the potential users of the evaluated system. By involving operators to verbalize their opinion, the symbiosis between operator and robot is promoted. Moreover, in the academic context, when creating a new teaching/research laboratory with collaborative robots, we are contributing to the synergy between disciplines, and to the promotion of creativity and innovation.

User research is the systematic study of the goals, needs and capabilities of users. Usability testing is a basic element for determining whether users are accomplishing their goals. Following the international standard definition ISO 9241, Part 11 [10], usability is the extend to which a product can be used by specific users to achieve specific goals with (i) effectiveness, (ii) efficiency, and (iii) satisfaction in a specified context of use. Effectiveness refers to the number of errors, number of successfully completed activities; efficiency relates with the task time, physical effort, fatigue, cognitive workload; finally, satisfaction is usually measured using subjective questionnaires. It is worth noting that other researchers [11] include usability in a broader methodological framework focused on user experience. A pleasurable user experience is an essential design target of human-collaborative robot interaction, and more goals could be added as: fellowship, sympathy, inspiration, and accomplishment.

The main objective in this research is to present a specific usability test plan to evaluate the usability of human-robot collaborative workspaces. It is developed through a concrete human-robot collaborative workspace experience (HRCWE) to illustrate how this usability test plan can be applied in a real environment. A description of the experience is provided, specifying objectives, roles, and responsibilities of all involved and associated timelines. It should be noted that main outcomes are not about the results obtained for this illustrative usability test plan, but the design of the plan itself.

This work is structured as follows. Section 2 introduces the purpose of the usability test in the HRCWE. Turn to this section to read the objectives of the test and understand why these objectives were selected. Next, it is described the experience under test and the context within which the experience is being working. Then, participants and responsibilities of everyone involved in the usability test are described, including testing team, observers, and test participants. Moreover, the evaluation procedure of the usability test is presented, including location and dates, test facilities and how test sessions are organized. Section 3 presents data collected during the usability sessions and contains the agreed usability metrics being tested. Next, results of experiment and statistics for evaluation and subsequent discussion are presented. A discussion of the findings of the experiment is provided in Section 4. Finally, Section 5 establishes some conclusions. The appendices contain useful working documents for the usability test, such as the recruitment screen, letters to participants, and proposed questionnaires.

2. The Proposed Usability Test

When a new collaborative robot is introduced into production, tasks under development change from the point of view of the human operator. It is necessary to evaluate how these changes affect human operator behavior, as well as measure and analyze human-robot collaboration tasks in detail.

2.1. Purpose of the Usability Test

The purpose of a usability test is, giving a specific context of use, measure the performance of a system in terms of task effectiveness, efficiency, and satisfaction. Thus,

the necessary feedback is provided to help decision-making in the redesign of systems. In particular, in this work, a collaborative workspace is considered where the operator is developing a main task. It is supposed that the operator have some experience in this main task and the workspace is correctly designed. Next, the operator is required for developing a secondary task, implying collaboration with a robot. The general objective of the usability test in this scenario would be to evaluate the usability of the workspace when a secondary collaborative task with cobots is added to the operator.

The usability of the proposed human-robot collaborative workspace's experience (HRCWE) is evaluated on the basis of the international usability standard ISO 9241-11 [10], which takes into account both objective and subjective metrics. According to this standard, the effectiveness and efficiency of usability is evaluated through the measurement of *Task to Time*, i.e., is, the time in seconds to complete a task, and *Task Completion rate* as the percentage of tasks that users complete correctly, performed by each participant while performing the tasks. In addition to these objective measures, a questionnaire has been developed to evaluate subjective usability using the System Usability Scale (SUS), whose definition is based on psychometric methods [12].

A usability test is a standard human-computer interaction test. However, human-robot interaction (HRI) is different from usual human-computer interaction (HCI) in several dimensions. HRI differs from HCI and human-machine interaction (HMI) because it concerns systems with complex, dynamic control procedures, exhibiting autonomy, cognition, and operating in changing real-world environments [13]. Our proposed usability test is not oriented to evaluate the design characteristics of the implemented prototype, nor its productive performance. Moreover, this test is focused in early steps of design, and not in final launch of products to the market. Human factors and ergonomics use too the same orientation in early steps of design. The aim is to use the collaborative robot as a partner, not as a substitute of human operator. Thus, the context implies the formation of human-robot teams, each contributing with their best skills [14].

2.2. Process under Experiment

Specifying the study's characteristics is the first step in a study of usability, to specify the context of use where the experimental study will be carried out, the description of the process, the type of participants, and the academic purpose of the workspace [15,16].

2.2.1. Context of Use

In this case, the experience is designed in a University laboratory, participants being recruited among students and teaching staff laboring in this environment. The laboratory endows two identical robot stations and a set of computer workplaces. It is designed to introduce students in the managing of collaborative robots in two training steps: the first one is understand the robot and how to program it, and the second step is adopting the role of human operator in a realistic scenario.

The human-robot collaborative workspace experience (HRCWE) is based on a prototype implemented in the laboratory with the aim of developing teaching and research on the relationships between humans and robots. In particular, in collaborative mode, with a focus on the cognitive and mental tasks of the human operator. Certainly, the use of a collaborative robot facilitates physical interaction with humans. However, the cognitive and mental aspects should not be underestimated. The perception of the human regarding the complexity of the task, or the trust that the human places in the robot are also relevant elements.

2.2.2. Procedure

To evaluate the effects on the human operator when a collaborative human-robot task is added in the original workspace, a workplace composed of two tasks is defined, in particular with a focus on assembly tasks. The workspace tasks are:

- Task 1: Tower of Hanoi, only performed by the human operator;

- Task 2: Collaborate with a cobot in the assembly of a product.

Task 1: Tower of Hanoi

The original task for the human operator is about solving the Tower of Hanoi (TOH5) with five pieces. This problem consists of five perforated disks of increasing radius stacked by inserting them into one of the three posts fixed to a board, as seen in Figure 1. The objective is to move one disk from the first pole to another one, making only one move at a time and placing a bigger disk on top of a smaller one. The Tower of Hanoi puzzle, was established as a robotics challenge as a part of EU Robotics coordination action in 2011 and IEEE IROS Conference in 2012. In our experiment, the Tower of Hanoi is performed only by the human operator.
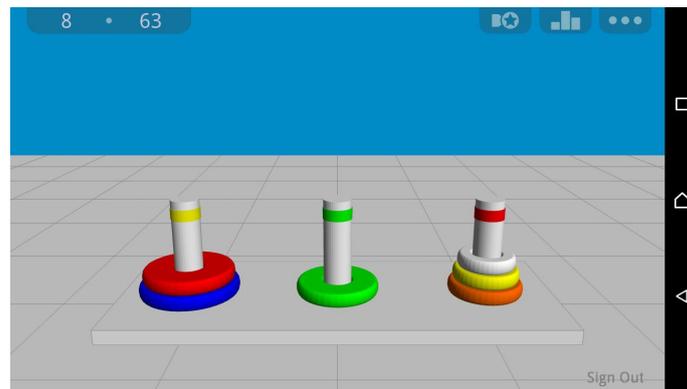
Figure 1. Digital version of the Tower of Hanoi problem with five disks, named TOH5.

A digital version of the TOH is used, available at Google Play HANOI 3D. This program allows manipulating the disks and records the number of moves and total time in seconds required to complete the task. No experimental cognitive variation exists if either wood pieces or a digital version is used [17].

Task 2: Collaborative assembly of a product

The introduced secondary task consists of collaborative assembling (CA) of a product composed of 3 components: a base, a bearing, and a cap, as shown in Figure 2. The task is classified as adding a low cognitive workload, from a human-centered perspective, because eye-hand coordinated skills are the more relevant in this task. Moreover, the assembly task shows a low level of physical risk for the human, and no action is necessary to decrease this risk.
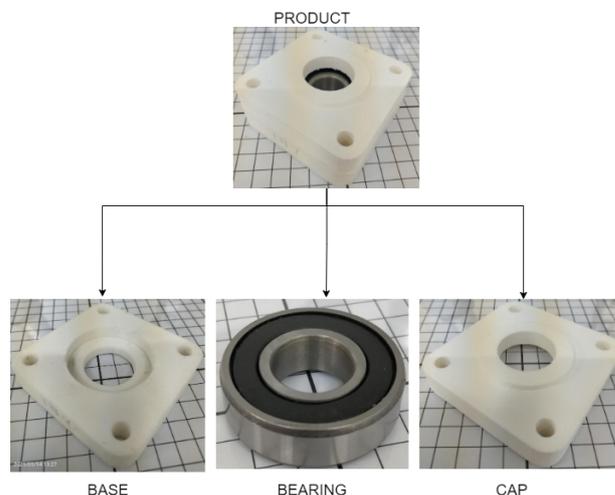
Figure 2. Collaborative assembly elements in the secondary process.

2.2.3. The Human-Robot Collaboration Workspace Experience, HRCWE

The HRCWE shown in Figure 3 is a workspace composed of two working areas: the work area one, called Tower of Hanoi with 5 disks (TOH5), is considered the main task to be developed by the human operator; the work area two, called Assembly, is a secondary added task where the collaboration with the cobot will be held (CA).
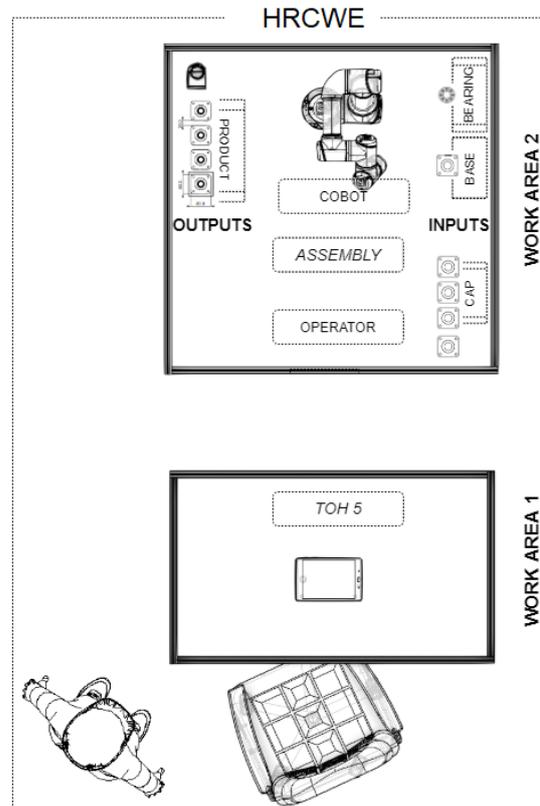
Figure 3. The human-robot collaboration workspace's experience, HRCWE.

In particular, the implemented assembly process area using a collaborative robot from the company Universal Robots, model UR3 is shown in Figure 4. The table where the cobot is anchored is divided into different work sub-areas. On the top right, a sub-area for parts feeding; on the bottom in the middle, a sub-area where the human executes actions on the teach pendant of the cobot; and on the top left, a sub-area where the human operator receives visual feedback of the cobot working, a red color light in a light tower.
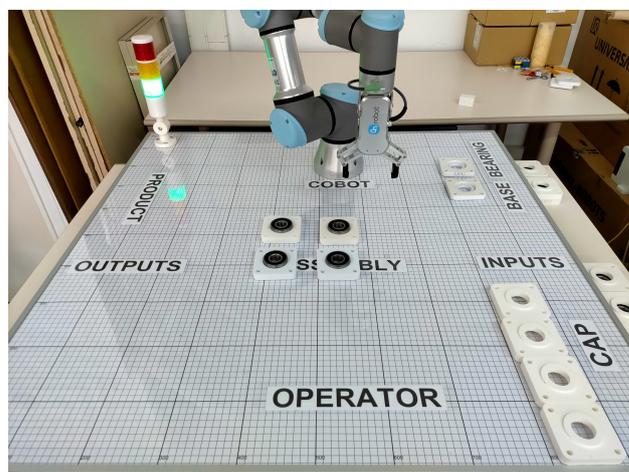
Figure 4. The collaborative assembly area implemented in the University laboratory.

2.3. Participants and Responsibilities

The roles involved in a usability test are as follows. It is worth noting that an individual may play multiple roles and tests may not require all roles.

2.3.1. Participants

Participants are University's bachelor students and some teaching staff. The participants' responsibilities are attempting to complete a set of representative task scenarios presented to them in as efficient and timely a manner as possible, and to provide feedback regarding the usability and acceptability of the experience. The participants are addressed to provide honest opinions regarding the usability of the application, and to participate in post-session subjective questionnaires. These participants have good skills in engineering methods, computer science and programming. They do not have previous knowledge about collaborative robots and how manage human-robot activities.

The form *Participants* in Appendix A contains a recruitment form to be used to recruit suitable participants. The more relevant elements in this form are:

- participant inclusion questions,
- participant exclusion questions, and
- participant experience questions.

2.3.2. Main Facilitator

General facilitator 's responsibilities are to:

- write the test plan,
- organize the recruitment of suitable participants,
- preserve ethical aspects,
- prepare the workspace for the development of the experimentation,
- show the task instructions to the user,
- record the data of the experiment,
- analyze usability test data, and
- summarize the results in a usability report.

The facilitator must supervise the ethical and psychological consequences for research participants. Any foreseeable effect on their psychological well-being, health, values, or dignity must be considered and, in the case of judging it negative, even to a minimal degree, eliminated [18]. From a robotic perspective, roboethics has as objective the development of technical tools that can be shared by different social groups. These tools aim to promote and encourage the development of robotics and to help in preventing its misuse against humankind [19].

The facilitator must ensure that the test can be carried out effectively. To do this, they must previously set the level of difficulty for the Tower of Hanoi solution—in this case 5 disks—, adjust the speed of the collaborative robot's movement, and program the cobot's task.

2.3.3. Ethics

All participants involved with the usability test are required to adhere to the following ethical guidelines:

- The performance of any test participant must not be individually attributable. The individual participant's name should not be used in reference outside the testing session.
- A description of the participant's performance should not be reported to his or her manager.

Considering that this study involves work with humans, the usability plan has been endorsed by the Ethics Committee of the UPC with the identification code 2021-06.

2.4. Evaluation Procedure

Elements for the evaluation procedure are now listed.

2.4.1. Location and Dates

The address of the test facility is: Automatic Control Department, FIB Faculty, Automatic Control Department Laboratory C5-202, Universitat Politècnica Catalunya Barcelona Tech. Authors plan to test participants according to the schedule shown in Table 1.

Table 1. Schedule of experiments.

Time	Monday	Wednesday	Thursday
09:00–11:00	Pilot Testing	Participant	Participant
11:00–13:00		Participant	Participant
14:00–16:00		Participant	Participant

2.4.2. Test Facilities

The experimental equipment under consideration is shown in Table 2.

Table 2. Experimental equipment.

Name	Description	Use
Tablet	Android system with digital version of TOH5 (HANOI3D)	Task 1 TOH5
Collaborative Robot	Robot Model UR3 with controller CB3. Manufactured by Universal Robot	Task 2 Collaborative Assembly (CA)
Laptop	Intel Core i5, Windows 10 Operating System	Data collection and logging
Web Cam	External, high definition	Get video of the experiment
Software	Cam video recorder, Visual Components v4.2	Record video of the experiment and record data of the collaborative robot

2.4.3. Pilot Testing

The purpose of the pilot test is to identify and reduce potential sources of error and fix any technical issue with the recording equipment or with the experiment that might cause delays to the current experiment. It is expected that the pilot test would take two hours, at maximum. The problems found would be immediately fixed. Observers are not invited due to the nature of the pilot.

2.4.4. Usability Sessions

Each participant session will be organized in the same way to facilitate consistency. Users will be interviewed at the end of the tasks.

Introduction

The main facilitator begins by emphasizing that the testing is being carried out by a Ph.D. student. This means that users can be critical of the experience without feeling that they are criticizing the designer. The main facilitator is not asking leading questions. The main facilitator explains that the purpose of the testing is to obtain measures of usability, such as effectiveness, efficiency, and user satisfaction, when working with collaborative robots. It is made clear that it is the system, not the user, that is being tested, so that, if they have trouble, it is the workspace's problem, not theirs.

Pre-Test Interview

At the beginning of the experiment, explained to the participant are: details of the experience, which are his/her tasks, how the experiment is developed (Appendix C is used for this), and that, at the end of the experiment, there is one questionnaire to be answered.

Participants will sign an informed consent that acknowledges: participation is voluntary, participation can cease at any time, and the session will be videotaped, but their privacy of identification will be safeguarded. The facilitator will ask the participant if they have any questions. The form *Consent form* in Appendix B is used for this aim. The more relevant aspects of this form are:

- description of objectives in the experiment,
- safety explanation for the participant, and
- participant's rights.

Next, the facilitator explains that the amount of time taken to complete the test task is measured and that exploratory behavior outside the task flow should not occur until after task completion. Time-on-task measurement begins when the participant starts the task.

The facilitator presents a demonstration to the user according to the guide *Demonstrations* in Appendix D. The more relevant aspects of this form are:

- demonstration of the use of the Tower of Hanoi game on the tablet;
- demonstration of operator involvement in product assembly

After all tasks are attempted, the participant completes the post-test satisfaction questionnaire.

Case Study

The experimental scenario is composed of two tasks within the HRWCE. The main task is Task 1 TOH5; a secondary task is added, Task 2 Collaborative Assembly (CA), as an additional human-robot collaboration task (see Figure 5). Table 3 shows the conditions of performance for the tasks. Time allocated for this scenario is 15 min.
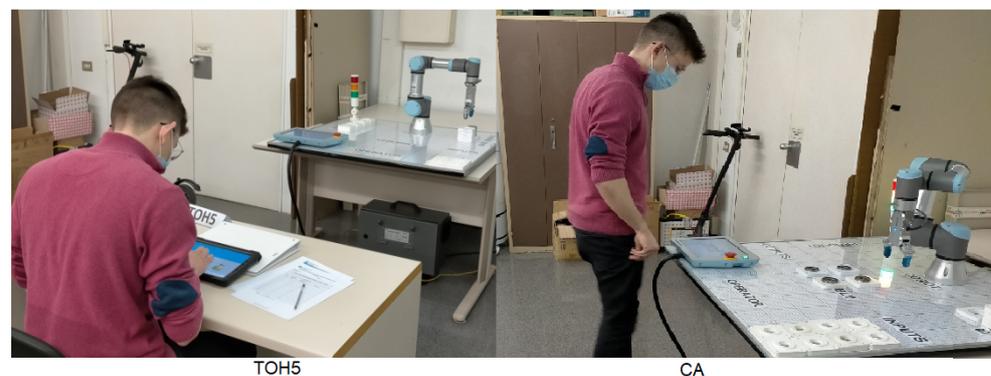
Figure 5. Scenario of the experience. Left, the TOH5 task, the main one, is performed. Right, the CA secondary collaborative assembly task is being developed.

Table 3. Experimental scenarios.

Task	Performance	Total Time
TOH5	Maximum number of TOH5 replays with 31 moves	15 min
CA	At least 7 work cycles completed	

The objective for the participant in the task TOH5 is to perform as many replays as possible. The secondary task, Collaborative Assembly, is related to responding to requests for collaboration from the cobot, which are indicated by the green light of the beacon in the assembly area. The time that the human takes to place caps is saved as *Wait Time* and

Cycle time and recorded in a data table, like the one shown in Table 4, jointly with figures for Task 1 when the operator is in the experimental scenario. In the collaborative assembly task, the activities of the participant, are:

- performing quality control of the assembly process,
- place the caps on the sub-assembly zone, and
- feeding the base and bearing warehouses.

Table 4. Form to Experimental Scenario (TOH5+CA). Tasks: Solve problem (main) and Collaborate with cobot (secondary).

Operator			Cobot		
Replay	**N_moves**	**Time to Task (s)**	**Cycle**	**Wait Time (s)**	**Cycle Time (s)**
1			1		
...			...		

Adapted Post-Test Questionnaire

At the end of the experience the participant answers the adapted System Usability Scale (SUS) as the satisfaction questionnaire shown in Table 5. As it has been shown [20], SUS can be applied to a wide range of technologies. This feature allows the questionnaire to be adapted to this particular experiment. In the SUS standard questionnaire, the word 'system' has been changed to 'human-robot collaborative workspace' because it is not a human-computer task but a human-robot task. A medium or low SUS score means that is necessary a discussion and effort redesign the experiment.

Table 5. Responses to individual statements of the SUS.

	Statements	Raw Score
1	I think that I would like to use Workspace Collab (HRCWE) frequently.	3.86
2	I found Workspace Human-Robot Collab (HRCWE) unnecessarily complex.	1.64
3	I thought Workspace Human-Robot Collab (HRCWE) was easy to use.	4.29
4	I think that I would need the support of a technical person to be able to use Workspace Human-Robot Collab (HRCWE).	1.57
5	I found the various functions in Workspace Human-Robot Collab (HRCWE) were well integrated.	4.0
6	I thought there was too much inconsistency in Workspace Human-Robot Collab (HRCWE).	2.07
7	I would imagine that most people would learn to use Workspace Human- Robot Collab (HRCWE) very quickly.	4.21
8	I found Workspace Human-Robot Collab (HRCWE) very cumbersome (awkward) to use.	1.50
9	I felt very confident using Workspace Human-Robot Collab (HRCWE).	4.29
10	I needed to learn a lot of things before I could get going with Workspace Human-Robot Collab (HRCWE).	1.43

3. Experimental Results

In a benchmark test, the usability of products is made measurable and enables a comparison with the competition. Based on different metrics, the usability dimensions, i.e., effectiveness, efficiency, and user satisfaction [10], are assessed and summarized into a meaningful overall score.

3.1. Data Collection and Metrics

A dataset with data collected from experiments is organized as shown in Figure 6. A set of statistical measures and tests available in the tool Usability Statistics Packages (Jeff Sauro's formulation is available online at: http://www.measuringusability.com/products/statsPak (accessed on 8 May 2021)) [21] are used for dataset analysis.
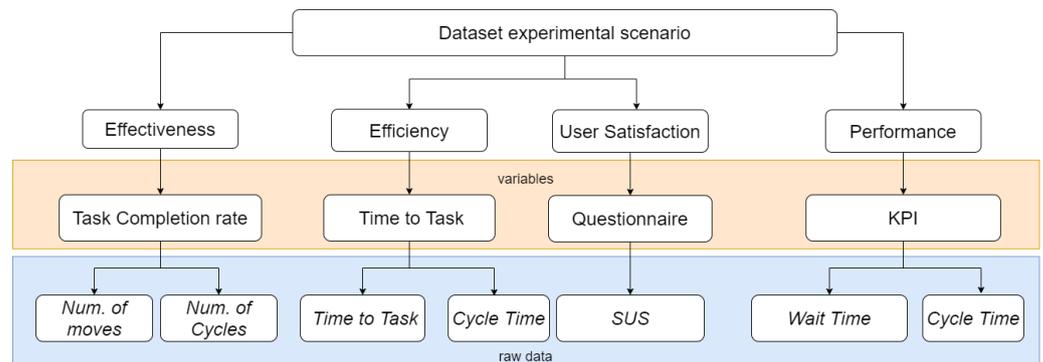
Figure 6. Organization of the dataset for the experimental study.

3.1.1. Task Effectiveness and Efficiency

The quantitative component involves handling of numerical variables and use of statistical techniques to guard against random events. This quantitative component includes information about the statistical significance of the results.

Effectiveness

The efficiency is evaluated using the *Task Completion rate* measure. For this binary variable, a maximum error of 10% of the optimal number of moves needed to solve the problem is set. Hence, it is coded with 1 (pass) for participants who solve the task and 0 (fail) for those who do not solve it.

Efficiency

To evaluate efficiency, the *Time to Task* measure is obtained from the dataset associated with TOH5. Only participants who completed the task are considered, and the analysis is made with the mean values obtained from each participation.

3.1.2. Satisfaction

The System Usability Scale (SUS) is used to evaluate the level of user satisfaction. The advantage of using SUS is that it is comparatively quick, easy, and inexpensive, whilst still being a reliable way of gauging usability. Moreover, the SUS questionnaire allows for providing us with a measure of people's subjective perceptions of the usability of the experience in the very short time available during evaluation sessions.

However, interpreting SUS scoring can be complex. The participant's scores for each question are converted to a new number, added together, and then multiplied by 2.5 to convert the original scores of 0–40 to 0–100. Though the scores are 0–100, these are not percentages and should be considered only in terms of their percentile ranking [22,23]. A way to interpret SUS score is to convert it into a grade or adjective [22], as shown in Figure 7.

Figure 7. Grade rankings of SUS scores adapted from Reference [22].

3.1.3. Key Performance Indicators for Cobot

To evaluate the CA task, i.e., the secondary task in the experimental scenario, some Key Performance Indicators (KPIs) have been collected, based in KPIs defined for cobots in Reference [24]. Table 6 shows the definitions used in this work.

The data gathering procedure to calculate cobot's KPIs has been implemented through a tool for obtaining the values of the variables recorded in the robot through the communication protocol with an external desktop. The values for *Cycle Time*, *Wait Time*, *Products* as the number of assembled products, and *Bases* as the number of bases, are acquired using this tool. Shown in Figure 8 is an example employing the Visual Components software, so this information is saved as an electronics sheet, for KPI analysis. The *CapTime* in the figure is the operator's time to place the cap. This value will be considered as the idle time of the robot, and also the *Human Operator Time*, since the robot is stopped.

Table 6. KPIs referred to cobots.

KPI	Definition
Cycle Time	Cycle Time measures the duration of one cobot sequence
Cycled Completed	How many cycles have been performed by the cobot in a particular time period
Per Utilization	How long a cobot is being used compared to how long it could
Per Efficiency	It defines the percentage of time that the cobot performs productive work while running a program
Wait Time	The percentage of time that the cobot is waiting while it is running a program
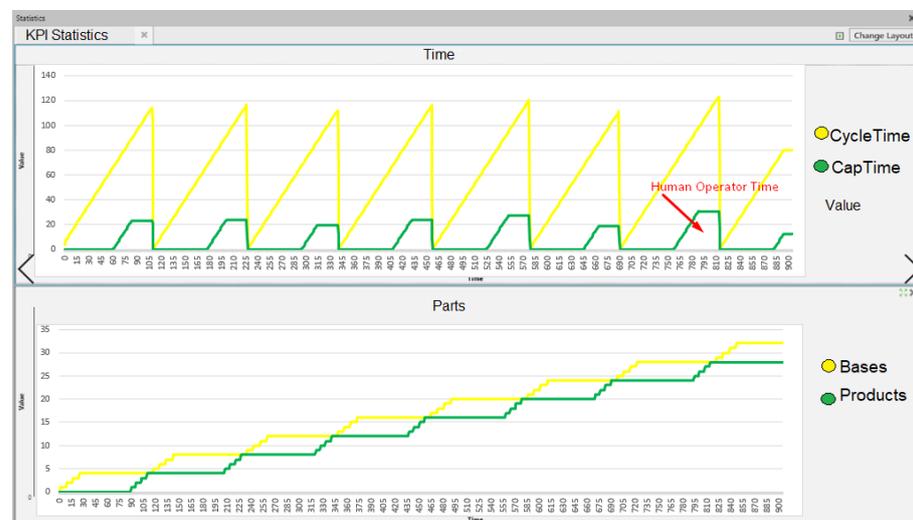
Figure 8. Time data and raw material sent from cobot UR3.

3.1.4. Video Recording

Experimental sessions are recorded on video; hence, the facilitator can later measure events, like: operator travel time from one task to another, how many times the operator check whether the robot has finished, and the right development of user actions in the assembly task.

While the user is developing the Tower of Hanoi task, the robot is working placing bases and bearings (red light). The vertical column of lights indicates with a green light when the presence of the user is required. The video recording can show if the user is paying attention to the column of lights or concentrating on the Tower of Hanoi task.

As a further analysis in collaborative human robot stations, video recording allows the observation of repetitive hand and arm movements that allow risk analysis and physical ergonomic assessment of the task.

3.2. Experimental Study

According to the usability test plan, the experiment dataset contains information for *Time to Task* and *Task Completion rate*, as well as results from questionnaire SUS and KPI values collected from 17 participants. Among them, 12 participants (70.6%) are undergraduate students, 2 participants (11.7%) are vocational students, and 3 participants (17.64%) from the teaching staff.

3.2.1. Results

Out of the seventeen participants, three are discarded as they have not correctly completed the proposed questionnaires. For this reason, results on fourteen participants are presented (n = 14). For the analysis, the following values are configured, the statistical significance level set at $p < 0.05$, and confidence level is 95%.

Effectiveness

The next 'pass and fail' histograms show the results in solving the tasks within the HRCWE: the histogram in Figure 9 shows 11 participants solving Task 1 (TOH5) and the histogram in Figure 10 shows the results of the human-robot (H-R) team in solving task 2. Values for *Task Completion rate* are calculated using the equation

$$Task\ Completion\ rate = \frac{Number\ of\ participants\ with\ successfully\ task}{Total\ number\ of\ participants\ undertaken}, \tag{1}$$

according to experimental results shown in Table 7.
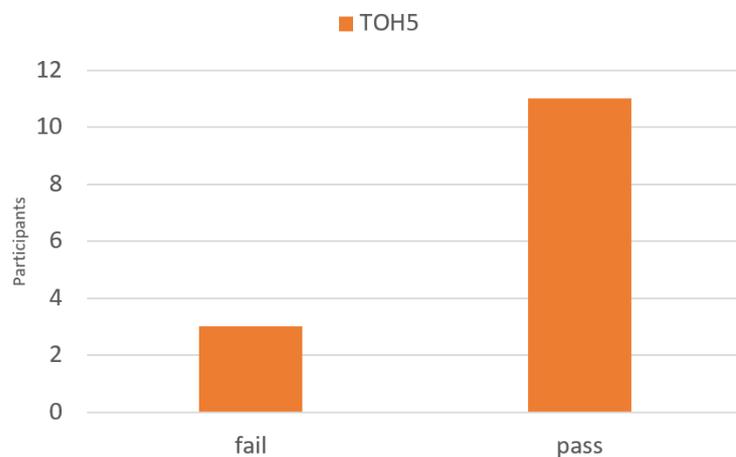
Figure 9. Histogram of fail and pass for tasks TOH5.
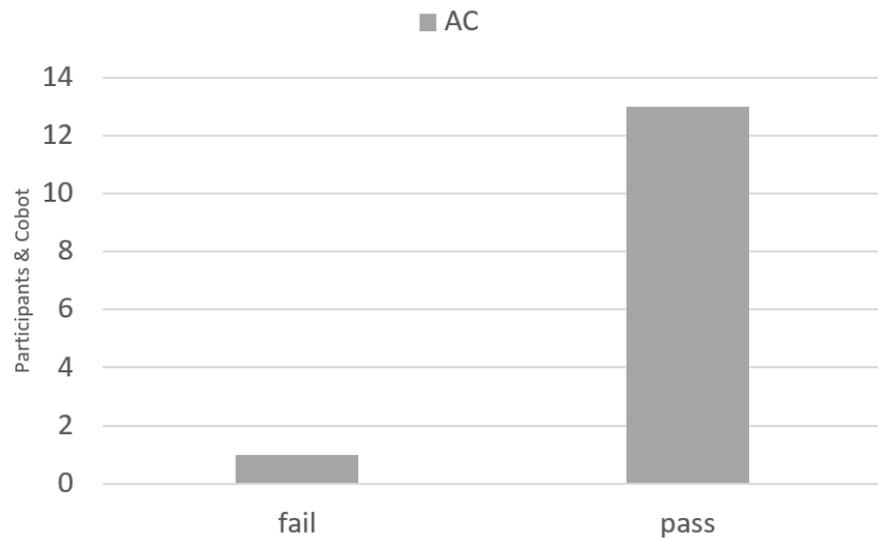
Figure 10. Histogram of fail and pass for tasks AC.

Table 7. Statistics *Task Completion rate*.

	TOH5		CA	
Succes	11		13	
n	14		14	
Task Completion rate	78.6 %		92.9%	
Confidence Interva	Low	High	Low	High
	51.70%	93.2%	66.5%	100%
Benchmark	56%			
p-value	0.044		0.002	

Sauro and Lewis experience states that a *Task Completion rate* lower than 56% could indicate a very low level of trustness for this variable. Results show a a superior level in both tasks (*p*-value = 0.044 and 0.002).

Results for the *Task Completion rate* level for TOH5 have a mean value of 78.6%. Based on benchmark with percentile Jeff Sauro, locate this value at percentile 50, and this value is acceptable as it is a first experience of the participant; however, for a continuous work, it is necessary to improve this value. The *Task Completion rate* level for the CA on 70% could be considered standard given the characteristics of human participation in the human-robot task.

Efficiency

To evaluate the efficiency, the *Time to Task* variable is analyzed. Firstly, with all the data obtained from the experiment, a percentile scale is generated, with five ranges for the variable defined, as shown in Table 8, according to the time spent in solving it.

Table 8. Percentiles in Time to Task to resolve TOH5 and CA.

Percentile	Time to Task (s)	
	TOH5	**CA**
10	23	102
50	44	107
75	66	113
90	93	116
98	>94	>120

　　　Task 1 (TOH5) is firstly analyzed. Statistical results for *Time to Task* are shown in Table 9, with a mean value of 56.2 s, brings the above percentile to 50%. Table 9 also shows values of the coefficient of variation (CV) (see Equation (2)) for the *Time to Task*. Task 1 (TOH5), with a value of 0.42, represents a high level of variation, as a result of solving the task only with human participation.

$$CV = \frac{\text{standard deviation}}{\text{mean value}}. \tag{2}$$

Table 9. Statistics of *Time to Task* for Task 1.

	TOH5	
Mean Value	56.2 (s)	
sd	25.9	
n	14	
Confidence Interval	Low	High
	43.34	72.89
CV	0.42	

　　　For Task 2, *Time to Task* is defined as the result of the team human-robot, as shown in Equation (3). This time is composed of the human's time (t_H) and the cobot's time (t_C).

$$\textit{Time to Task} = t_H + t_C. \tag{3}$$

Figure 11 shows the results obtained by each participant. It can be observed the time composition of the task.

　　　For the statistical analysis, we consider *Time to Task* equal to *Cycle Time* for the cobot, and t_H equal to *Wait time*, obtained directly from the cobot controller.

　　　Results in Table 10 show a *Time to Task* of 108.9 s as mean value. This corresponds to a percentile of 50%. CV value is 0.04, equivalent to 4% of variation. This is considered a low variation. For t_H, the mean value is 17.05 s, with CV 0.25 equivalent to 25%. In this case, this is considered a high variation, typical for human tasks. Finally, for t_C, the mean value is 90.58 s, and CV is down to 0.027, equivalent to 2.7%, which is a minimum variation, as expected, considering the high stability of the cobot.
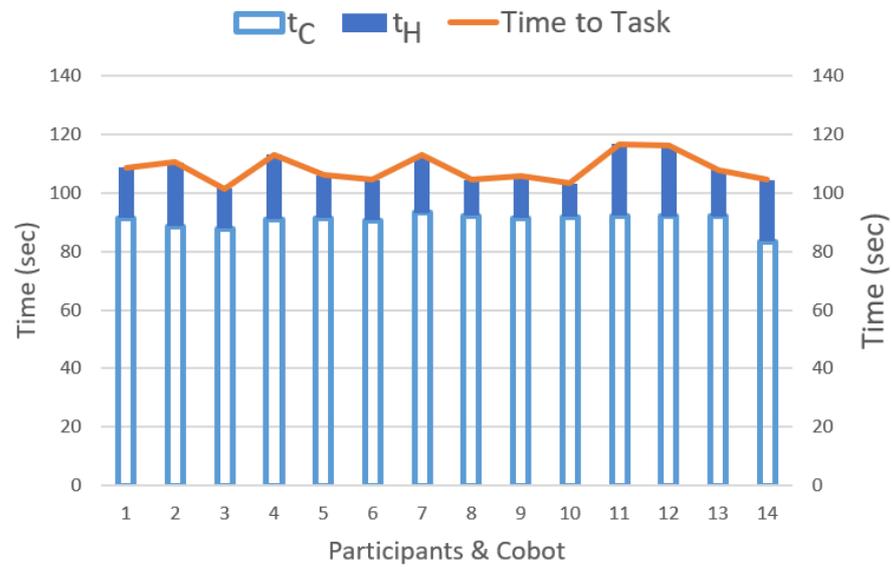
Figure 11. *Time to Task* for Task 2.

Table 10. Statistics times in Task 2.

	Time to Task (s)	t_H s	t_R s
Mean value	108.9	17.05	90.58
sd	4.8	4.5	2.5
n	14	14	14
Confidence Interval			
Low	105.36	14.7	89.10
High	110.89	19.79	92.08
CV	0.04	0.25	0.027

3.2.2. System Usability Scale Score Results

Table 11 provides an overview of mean values, standard deviations, incomplete questionnaires, and checking the coding of the questionnaires, as well as reliability indices values (Cronbach's α) for the obtained measures.

Table 11. Statistics SUS scoring and reliability test.

Mean value	81.1		
sd	13.3		
Non-Blank	14	Coding Check	Values appear to be coded correctly from 1 to 5
Cronbach Alpha	0.814	Internal Reliability	Good

Table 12 shows results about SUS, the interpretation in percentiles, and a descriptive adjective of its value. The value of 81.1 qualifies the HRCWE as *Excellent* and the degree of acceptability as *Acceptable*.

Table 12. SUS results interpretation.

Raw SUS score	81.1	Percentile Rank	88.1%
SUS Benchmark	Hardware	Adjective	Excellent
		Grade (Bangor)	B
		Grade (Sauro & Lewis)	A-
		Acceptability	Acceptable

Considering HRCWE as a hardware system, following the Sauro and Lewis classification, the benchmark of experience with hardware shows that a raw SUS score of 81.1 has a higher SUS score than 88.14% for Hardware.

The main value from SUS is providing the single total score. However, it is still necessary to look into detail the individual score for each statement [22]. This information in presented in Table 5. Caglarca [25] suggested taking into account individual evaluation by verifying the shape of a "five-pointed star" visualization. Hence, raw scores from Table 5 have been transformed into a radial chart, as shown in Figure 12. Caglarca also concluded that the more the five-pointed star looks, the more positive the usability will get. Although it tends to be a subjective assessment, it is worth noting that the five-pointed star shape is almost in perfect form, as in this study.
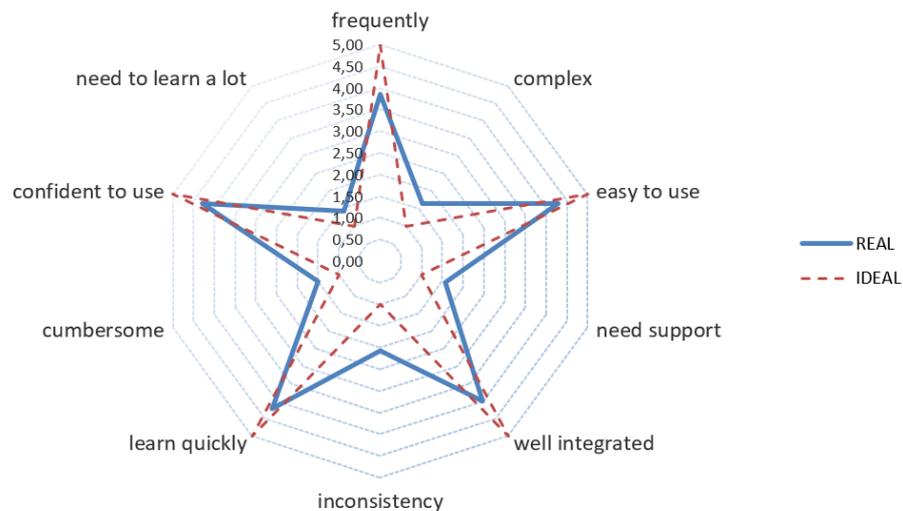
Figure 12. Evaluation of the five-pointed star.

3.2.3. Performance of Cobot

To finalize the experiment, with the data of *Cycle Time* and *Wait Time*, the KPIs of Per Efficiency and Per Utilization of the cobot are obtained to evaluate the performance of the Cobot in HWRCE. The Per Utilization is calculated as,

$$Per\ Utilizacion(participant) = \frac{\overline{Cycle\ Time} - \overline{Wait\ Time}}{\overline{Cycle\ Time}} * 100. \tag{4}$$

The statistical analysis in Table 13, with a mean value of 83%, shows an Per Utilization greater than 80% in the use of cobot for task collaborative assembly.

Table 13. Statistics of *Per_Utilization* for the cobot in HWRCE.

Mean Value	83%		
sd	0.03		
n	14		
Confidence Interval	Low	High	
	81%	85%	
Benchmark	80%	*p*-value	0.008

The percentage value of the efficiency (Per Efficiency) is calculated considering the total time of a work cycle; in this case, it was set at 900 s.

$$Per\ Efficiency = \overline{NumberCyclesCompleted} * \frac{\overline{Cycle\ Time}}{900} * 100. \tag{5}$$

The statistical analysis of the Per Efficiency in Table 14 shows with a mean value of 84% a Per Efficiency higher than the 75% (*p*-value = 0.03) established as a reference in this experiment.

Table 14. Statistics *Per Efficiency* of cobot in HWRCE.

Mean Value	84%		
sd	0.13		
n	14		
Confidence Interval	Low	High	
	74%	96%	
Benchmark	75%	*p*-value	0.03

4. Discussion

The feasibility of extrapolating the usability experience in HCI towards HRI is clearly defined along with this study: the context of use, requirements, workspace design, task allocation between human and robot, experimental testing, and validation steps.

In this study, with a *Task Completion rate* value of 78.6 in the effectiveness of the task 1, it can be considered that the human operator can effectively solve Task 1 in the HRCWE. To increase the effectiveness in this task, a first alternative is the incorporation of a training stage, and a second alternative could be the use of an assistant that considers assisting the operator when the real-time value of the *Task Completion rate* is lower than a minimum set value. For the second task, the value of the *Task Completion rate* shows that the human-robot team effectively solves the Task 2. However, a redesign of the physical architecture of HRWCE, in which the human operator is closer to the work area, could improve the efficiency of the work team.

The efficiency in Task 1, measured through the *Time to Task* with mean values of 56.2 s for Task 1 and 108.09 s for Task 2, places the efficiency between the low and standard level, with a higher accumulation towards the standard level. Hence, it can be considered that the human operator is able to efficiently solve the tasks in the HRCWE.

The SUS score shows that the collaborative workspace is perceived as acceptable for working with humans, and the star chart proves that the performance of its components is balanced, as expected.

The evaluation carried out about HWRCE, through the KPIs, corroborates the capacity of the team human-robot, with values higher than 80% of Per Utilization and higher than 75% of Per Efficiency, and a value of *Task Completion rate* over 80%. The variability analysis show that the system is able to absorb the variability introduced by the human operator.

In order to improve the results obtained in the efficiency and effectiveness of the tasks within the HWRCE, adding the real-time variable *Difficulty Task*, which considers the variables *Task Completion rate* and *Time to Task*, could be used to work with an assistant to guide a strategy for solving the tasks.

Overall, the usability benchmark additionally demonstrates the flexibility of the human operator to work in conjunction with a cobot operator in collaborative assembly tasks within the HRCWE.

5. Conclusions

This article introduces a methodological and systematic guide for experimenting and evaluating experiments related to human-robot interaction in assembly task workspace. Taking advantage of usability in human-computer interaction, this experience has been expanded and adapted to the field of collaborative human-robot interaction to have a solid and well-founded basis for evaluating the collaborative workspace, where the human operator shares tasks with a robot. Reviewing and incorporating best practices from relations area can reduce the number of testing iterations required and save time and money developing and evaluating a process or system.

In the future, it is expected to expand this guide for the assembly of products with a greater number of components and/or with variants in components. The following experimentation will evaluate the different forms of human-robot collaboration with other architectural of workspace. It is also expected to use it for other types of human-robot collaborative work applications, such as picking, packing, and palletizing.

If usability is included within the user experience, there is research work to be done. It is convenient to expand the registry of variables to be measured and take into account aspects of trust, acceptance of technology, and empathy [26].

Author Contributions: Conceptualization, A.C. and P.P.; methodology, A.C. and P.P.; validation, A.C.; formal analysis, P.P.; investigation, A.C. and P.P.; resources, C.A.; writing—original draft preparation, A.C., P.P. and C.A.; writing—review and editing, A.C., P.P. and C.A.; supervision, C.A.; project administration, C.A.; funding acquisition, P.P. and C.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been co-financed by the European Regional Development Fund of the European Union in the framework of the ERDF Operational Program of Catalonia 2014–2020, grant number 001-P-001643. Prof. Cecilio Angulo has been partly supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 825619 (AI4EU). Prof. Pere Ponsa has been partly supported by the Project EXPLainable Artificial INtelligence systems for health and well-beING (EXPLAINING) (PID2019-104829RA-I00/AEI/10.13039/501100011033).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Participants Selection

We need 30 participants in total. Each session lasts <25> min. We must have a candidate's completed screener at least one day in advance of the session, so we know which experimental group to assign him or her to.

- Introduction:
  This experiment is part of a research project related to Human Robot Interaction (HRI), the basic objective is to determine the variation of mental load on the operator when a collaborative task with a robot is added.
- Selection questions:
  1. The participant is familiar with information and communication technologies.

2. The participant is interested in the use of robotics and its applications.
3. Participant feels confident working with a moving robot.
4. Participant would like to help in this research.

- Exclusion questions:
  1. The participant is of legal age
  2. The participant feels insecure working with automatic machines.
- Experience Questions
  1. Do you have experience in programming or using robots?
  2. Have you participated in projects related to mind uploading?
  3. Do you have experience playing the towers of Hanoi, in real physical format or in its digital version?

Appendix B. Statement of Informed Consent. Consent Form

TITLE: Usability Test in the Human Robot Collaborative Workspace

PROTOCOL DIRECTOR: Phd. student Alejandro Chacón.

DESCRIPTION: You have been invited to participate in a study that aims to improve the tasks performed by operators and cobots in workspaces within factories.
The facilitator gives you the instruction (Development of the game Hanoi's Tower and the collaboration with the robot in the assembly of the product)

TIME INVOLVEMENT: Your participation will take approximately 15 min

RISKS AND BENEFITS: There aren't risks in this study. The benefits are only for academic purposes. Your decision whether or not to participate in this study will not affect your grades in school.

PAYMENTS: You will not receive anything as payment for your participation. In fact, you will receive feedback about the experimental session.

SUBJECT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. The alternative is not to participate. You have the right to refuse to answer particular questions. Your individual privacy will be maintained in all published and written data resulting from the study.

CONTACT INFORMATION:
Questions: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, contact the Protocol Director, Alejandro Chacón, luis.alejandro.chacon@upc.edu.
I give consent for my identity to be revealed in written materials resulting from this study only inside the class with my teacher and colleagues:
Please initial: _Yes _ No

The extra copy of this consent form is for you to keep.

SIGNATURE   DATE

Appendix C. Case Study

Two different task are defined on these scenario, with different conditions and operating characteristics, as shown in Table A1. Each participant participates in two tasks, being always the task1 the first one. One iteration of scenario is performed for each operator for 15′.

Table A1. Experimental scenario.

	Task	**Performance**	**Time**
1	TOH5	Maximum number of TOH5 game replays with 31 moves	
2	CA	At least 7 Cycles Work completed	15 min

The objective for the participant in the TOH5 throuble is to perform as many repetitions as possible. The number of movements and the time of each repetition are recorded by the participant in a data table as Table A2.

The second task, Assembly, is related to responding to requests for collaboration from the robot, which are indicated by the green light of the beacon in the area of assembly. The time that the human takes to place caps, defined as Wait Time and Cycle time, are recorded in a data table as Table A2, jointly with figures for Task 1 when the operator is in the scenario 2. In the assembly task, the activities of the participant, are:

- performing quality control of the assembly process,
- place the caps on the sub-assembly zone, and
- refill the base and bearing warehouses.

Table A2. Scenario. Task 1 & Task 2: Resolve TOH5 & Collaborate with cobot (CA).

Operator			**Cobot**		
Replay	**N_moves**	**Time to Task (s)**	**Cycle**	**Wait Time (s)**	**Cycle Time (s)**
1			1		
…			…		

At the end of the experiment the participant answers, the System Usability Scale (SUS) as a satisfaction questionnaire.

Appendix D. Demonstrations

The main facilitator shows the participant the two areas and how the tasks are performed, in particular highlighting the activities that the operator must perform.

Appendix D.1. TOH5

By using the app's own functions, the facilitator shows once how to solve the game with the least number of moves; see Figure A1.
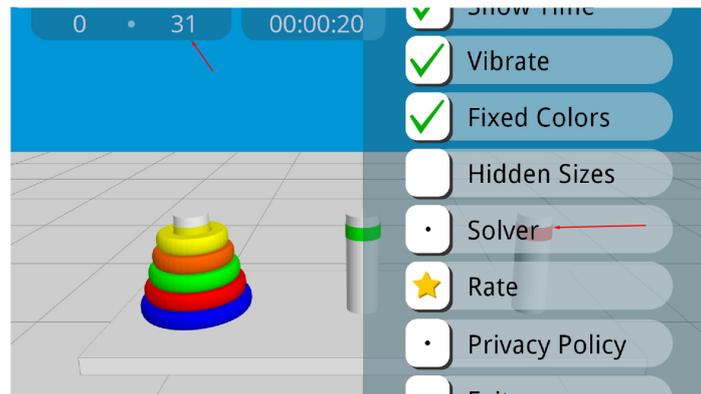
Figure A1. TOH5–Solver.

Appendix D.2. Assembly

The facilitator shows a complete work cycle, indicating the activities that the operator must perform: place the caps, click on the teach pendant, and reload stores, as well as the meaning of the lights on the indicator tower; see Figure A2.
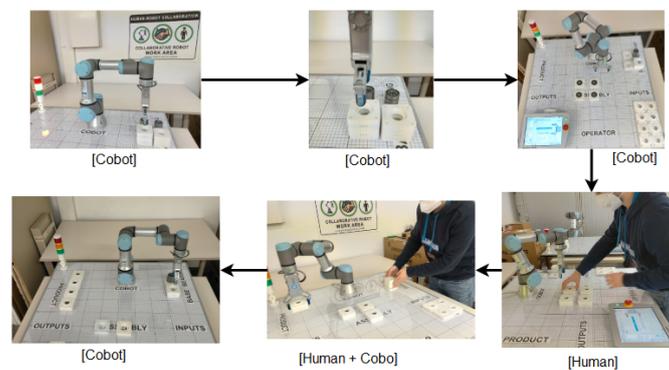
Figure A2. Demonstration assembly cycle work human-hobot.

References

1. Nahavandi, S. Industry 5.0—A Human-Centric Solution. *Sustainability* **2019**, *11*, 4371. [CrossRef]
2. International Organization for Standardization. *TS 15066: 2016: Robots and Robotic Devices–Collaborative Robots*; Standard; International Organization for Standardization: Geneva, CH, Switzerland, 2016.
3. ISO Central Secretary. *Ergonomics of Human-System Interaction—Usability Methods Supporting Human-Centred Design*; Standard; International Organization for Standardization: Geneva, CH, Switzerland, 2002.
4. Maurtua, I.; Ibarguren, A.; Kildal, J.; Susperregi, L.; Sierra, B. Human–robot collaboration in industrial applications: Safety, interaction and trust. *Int. J. Adv. Robot. Syst.* **2017**, *14*. [CrossRef]
5. Davis, F.D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q.* **1989**, *13*, 319–340. [CrossRef]
6. Gervasi, R.; Mastrogiacomo, L.; Franceschini, F. A conceptual framework to evaluate human-robot collaboration. *Int. J. Adv. Manuf. Technol.* **2020**, *108*, 841–865. [CrossRef]
7. Vanderborght, B. *Unlocking the Potential of Industrial Human-Robot Collaboration*; Publications Office of the European Union: Luxembourg, 2020. [CrossRef]
8. Pacaux-Lemoine, M.P.; Berdal, Q.; Guérin, C.; Rauffet, P.; Chauvin, C. Designing human–system cooperation in industry 4.0 with cognitive work analysis: A first evaluation. *Cogn. Technol. Work.* **2021**. [CrossRef]
9. Marvel, J.A.; Bagchi, S.; Zimmerman, M.; Antonishek, B. Towards Effective Interface Designs for Collaborative HRI in Manufacturing: Metrics and Measures. *J. Hum.-Robot Interact.* **2020**, *9*. [CrossRef]
10. ISO Central Secretary. *Ergonomics of Human-System Interaction. Usability: Definitions and Concepts*; Standard ISO 9241-11:2018; International Organization for Standardization: Geneva, CH, Switzerland, 2018.
11. Chowdhury, A.; Ahtinen, A.; Pieters, R.; Vaananen, K., User Experience Goals for Designing Industrial Human-Cobot Collaboration: A Case Study of Franka Panda Robot. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*; Association for Computing Machinery: New York, NY, USA, 2020.

12. Lewis, J.R. *Usability Testing, Handbook of Human Facto rs and Ergonomics*; John & Wiley: Hoboken, NJ, USA, 2006; pp. 1275–1316.

13. Scholtz, J. Theory and evaluation of human robot interactions. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 6–9 January 2003. [CrossRef]

14. Tekniker, Pilz and ZEMA. *Definition and Guidelines for Collaborative Workspaces*; Technical Report GA Number 637095; European Commission: Brussels, Belgium , 2017.

15. Kalmbach, S.; Bargmann, D.; Lindblom, J.; Wang, W.; Wang, V. Symbiotic Human-Robot Collaboration for Safe and Dynamic Multimodal Manufacturing Systems. 2018. Available online: https://cordis.europa.eu/programme/id/H2020_FoF-06-2014 (accessed on 12 February 2021).

16. Masó, B.; Ponsa, P.; Tornil, S. Diseño de tareas persona-robot en el ámbito académico. *Interacción Rev. Digit. De AIPO* **2020**, *2*, 26–38.

17. Hardy, D.J.; Wright, M.J. Assessing workload in neuropsychology: An illustration with the Tower of Hanoi test. *J. Clin. Exp. Neuropsychol.* **2018**, *40*, 1022–1029. [CrossRef] [PubMed]

18. ETSI. *ETSI Guide: Human Factors (HF); Usability Evaluation for the Design of Telecommunication Systems, Services and Terminals*; Standard ETSI EG 201472 2000; ETSI: Sophia Antipolis, France, 2000.

19. Veruggio, G.; Operto, F., Roboethics: Social and Ethical Implications of Robotics. In *Springer Handbook of Robotics*; Siciliano, B., Khatib, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1499–1524. [CrossRef]

20. Cowley, A.W. IUPS—A retrospective. *Physiologist* **2006**, *49*, 171–173. [PubMed]

21. Zazelenchuk, T.; Sortland, K.; Genov, A.; Sazegari, S.; Keavney, M. Using Participants' Real Data in Usability Testing: Lessons Learned. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*; CHI EA '08; Association for Computing Machinery: New York, NY, USA, 2008; pp. 2229–2236. [CrossRef]

22. Bangor, A.; Kortum, P.T.; Miller, J.T. An empirical evaluation of the system usability scale. *Int. J. Hum.-Comput. Interact.* **2008**, *24*, 574–594. [CrossRef]

23. Sauro, J.; Lewis, J.R. *Quantifying the User Experience. Practical Statistics for User Research*; MK Morgan Kaufmann: Burlington, MA, USA, 2016.

24. Bouchard, S. *Lean Robotics: A Guide to Making Robots Work in Your Factory*; Amazon Italia Logistica: Torraza Piemonte, Italy, 2017; p. 222.

25. Wang, Y. System Usability Scale: A Quick and Efficient User Study Methodology. 2018. Available online: http://ixd.prattsi.org/2018/04/system-usability-scale-a-quick-and-efficient-user-study-methodology/ (accessed on 6 March 2021).

26. Prati, E.; Peruzzini, M.; Pellicciari, M.; Raffaeli, R. How to include User eXperience in the design of Human-Robot Interaction. *Robot. Comput.-Integr. Manuf.* **2021**, *68*, 102072. [CrossRef]