Assessing perceived assembly complexity in Human-Robot collaboration processes: a proposal based on Thurstone's Law of Comparative Judgement

Matteo Capponi[1], Riccardo Gervasi[2], Luca Mastrogiacomo[3] and Fiorenzo Franceschini[4]*

[1] *matteo.capponi@polito.it {ORCID: 0000-0002-4669-2140}*
[2] *riccardo.gervasi@polito.it {ORCID: 0000-0002-3006-3382}*
[3] *luca.mastrogiacomo@polito.it {ORCID: 0000-0002-8454-5918}*
[4] *fiorenzo.franceschini@polito.it {ORCID: 0000-0001-7131-4419}*

Politecnico di Torino, DIGEP (Department of Management and Production Engineering),
Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

*Corresponding Author

Abstract

Due to the growing demand for customized products, companies have faced increasing product and process complexity levels. To address this issue, manufacturing processes should become more flexible. One of the most promising technologies to achieve this goal is collaborative robotics (or "cobots"). In collaborative assembly processes, man and robot cooperate: cobots perform heavier tasks, while humans intervene when more dexterity is required. However, the co-existence of humans and cobots in the same workspace may influence the operators' perception of assembly complexity. The analysis and control of assembly complexity are crucial to achieving better performances in terms of process quality and operators' well-being. Many qualitative methods have been proposed in the literature to provide a holistic assessment of assembly complexity. This paper proposes a novel method to define a quantitative scale of perceived assembly complexity, based on Thurstone Law of Comparative Judgements. This method was applied to an experimental case-study concerning the assembly of three different products in two modalities (i.e., manual and collaborative). Regression analysis showed that the perceived complexity may be related to the occurrence of process failures and to the perceived workload. The method also proved capable of identifying assembly processes where cobot assistance was helpful, providing process designers with a supporting tool to minimise perceived complexity.

Keywords: Perceived assembly complexity, product complexity, Thurstone Law of comparative judgement, Human-robot collaboration, process failures, perceived workload.

Paper type: Research paper

1. Introduction

Recent years have seen a renewed interest for humans in manufacturing, resulting in the new Industry 5.0. paradigm. This concept promotes an anthropocentric view of production in which humans assume a primary role. In such a context, technology is not adopted to replace humans, but rather to improve their working conditions and well-being. One of the most promising technologies to address this challenge is collaborative robotics, i.e., special robots ("cobots") that can work closely with humans, sharing workspace and goals (Maddikunta *et al.* 2022). Space sharing, which was forbidden with traditional robots for safety reasons, represents the great contribution of such technology and, simultaneously, one of the biggest concerns (Zanchettin *et al.* 2016, Villani *et al.* 2018, Gualtieri *et al.* 2021). Therefore, cobots can be used for repetitive and strenuous tasks, while humans can perform higher value-added activities, thanks to their flexibility and dexterity. This creates a synergetic working environment based on the collaboration between human and machine, the so-called "human-robot collaboration" (or "HRC") (Bauer *et al.* 2008). Nonetheless, to design an efficient human-robot collaboration a wide variety of aspects should be taken into account, e.g., robot's adaptivity, the means of communication, task organization, the psychophysiological impact on humans, the fluency of collaboration, etc. (Hoffman 2019, Gervasi *et al.* 2020, Sigurjónsson *et al.* 2022).

In manufacturing environments, cobots are becoming increasingly common, especially in assembly processes, However, few studies investigated the effects of using cobots in assemblies in terms of process quality. To this end, the authors proposed to consider as driving concept that of "assembly complexity", which has proven to be correlated with the effort required to complete an assembly process, process quality and costs (ElMaraghy *et al.* 2012, Alkan 2019, Verna *et al.* 2022a).

Generally, given the large variety of dimensions to be considered, the use of qualitative approaches for complexity assessment based on surveys and questionnaires is unavoidable, even in the case of manual assemblies. However, such methodologies present some limitations regarding the qualitative nature of the results they provide. In this framework, the authors proposed a novel "TICS method" (i.e., Thurstone-Inspired Complexity Scaling method), to assess perceived assembly complexity using Thurstone's law of comparative judgements (Thurstone 1927). The method was implemented to assess complexity for different products and different assembly modalities (i.e., collaborative and manual). Two effects of perceived assembly complexity were addressed: process failures (objective) and perceived workload (through NASA-TLX tool (Hart and Staveland 1988)). The paper is organised as follows: section 2 provides a state-of-the art on assembly complexity in manual and collaborative processes; section 3 introduces TICS method. Section 4 describes the implemented methodology and section 5 provides the respective results. Section 6 summarises the main findings, limits and potential future developments of this work.

2. Literature review

In this section the literature review will be addressed from two different perspectives: assembly complexity in both manual and in collaborative assembly processes.

2.1 Assembly complexity

The term "complexity" is a very broad concept that encompasses a wide variety of variables (ElMaraghy *et al.* 2012, Alkan *et al.* 2018). In manufacturing, assembly complexity is mainly concerned with the peculiarities of products, processes and production systems that companies have to manage. This may result in a higher degree of skills required from workers, greater costs and a deterioration of product and process quality (Hvam *et al.* 2020). Therefore, a model able to measure complexity can support product and process designers to take cost-effective preventive and corrective actions.

Regarding manual assembly, there is a clear distinction between objective assembly complexity and perceived assembly complexity (Falck, Örtengren, *et al.* 2017a, Alkan 2019):

- Objective complexity represents an intrinsic property of the assembly process, and it is independent on the subject performing the task. It is related to factors like the number and variety of elements involved (e.g., components and connectors), interaction and dependences between elements, assembly sequences, components' geometrical features, etc.
- Perceived complexity refers to the subjective experience of complexity within an assembly process. It is strongly related both to objective complexity and depends on the personal capabilities and experience of the observer or performer involved in the process.

2.1.1 Objective assembly complexity

Over the years, several quantitative models have been developed to objectively assess assembly complexity. One major stream of research relates assembly complexity specifically to product complexity, which refers to all dimensional, geometrical, and structural characteristics of a certain product. Many of this models derived from Design For Assembly ("DFA") principles (Boothroyd and Alting 1992, Boothroyd 1994, Battaïa *et al.* 2018). Due to the spread of industrial robots in manufacturing, the so-called "DFA2" methodology (i.e., "Design for Automatic Assembly) was also introduced. It consists of a set of design rules for products whose assembly process is fully automated (Eskilander 2001, Roulet-Dubonnet *et al.* 2018, Madappilly and Mork 2021). Regarding, objective assembly complexity, a pioneering work in this field was proposed by Hinckley (1994) who defined a complexity factor basing on assembly times and underlined the importance of reducing complexity to enhance process quality and costs. Similarly, Shibata first (2002) and later Su et al. (2010) linked assembly complexity to two factors: i.e., the "design-based complexity factor" derived from geometrical and dimensional evaluations of products' features, and the "process-based complexity factor" calculated using standard assembly times. Time, indeed, is often used as an indirect measure of assembly complexity, since the greater time is needed to assemble a product the greater its complexity. In this context, Alkan (Alkan 2019) theorized a novel method to assess assembly complexity based on assembly standard times and DFA theory. This method exploited a more generalized product complexity model developed by Sinha (Sinha 2014, Sinha and de Weck 2014). In Sinha's model products are assimilated to molecular structures and their complexity depends on individual component complexity, interface complexity and topological complexity. Similarly, Verna et al. (2022) modified Alkan's complexity model and used it to predict defects of assembled products. Also Sudhoff et al. (2022) proved the existence of relationship between complexity measures and assembly times.

Another common approach to evaluate assembly complexity consists of applying principles of information theory (Shannon 1948) to products, production processes and systems. These methods rely on the assumption that complexity and difficulties emerge when uncertainty is involved in the assembly process. They adapted the concept of information entropy, that is a measure of uncertainty of a random signal (Shannon 1948), to assembly processes. Uncertainty may depend on variety of components or fasteners, on assembly sequences, on tools and on products demanded, etc. ElMaraghy and Urbanic (2003, 2004) proposed a novel entropy-inspired method (called "MCAT" , i.e., Manufacturing Complexity Assessment Tool) that relates manufacturing complexity to quantity, diversity and content of information to be managed (Capponi *et al.* 2023). Fujimoto et al. (2003) used information entropy to develop a methodology to manage manufacturing complexities of assembly systems due to product varieties. Similarly, Zhu et al. (2008) considered product varieties as the main source of manufacturing complexity and introduced an entropy-based complexity measure called "operator choice complexity". It refers to difficulties that arise when various choices must be made by the operators facing a wide variety of products to assemble. Ameri et al. (Ameri *et al.* 2008) combined information theory and graph theory to provide a model to assess product design complexity. Subsequently, ElMaraghy W. and Urbanic's model (ElMaraghy and Urbanic 2003, 2004) was modified and combined with DFA principles to assess product assembly complexity (Samy and

ElMaraghy 2010). A similar model is further used by Samy and ElMaraghy H. (2012) to develop a metric of complexity suitable for the whole manufacturing system. This metric can be used by designers to reduce assembly costs and improve quality. Wang and Hu (2010) developed a complexity measure based on the uncertainty of operator's choice in assembly systems with different configurations (e.g., parallel or hybrid), taking into account also operator reaction times and fatigue. The model was subsequently used to reduce manufacturing complexity in mixed-model assembly systems (Wang *et al.* 2013). Similarly, to level manufacturing complexity Zeltzer et al. (2017) defined an entropy-based complexity measure that takes into account the variability of task duration in mixed assembly lines. Sun and Fan (2018) introduced the concept of "changeover complexity" that refers to difficulties and uncertainty perceived by operators during the assembly of products with multiple option features. Even in this case, the choice among different parts, tools, fasteners, etc. increases perceived complexity that is thus quantified using information entropy. More recently, Liu et al. (2021) developed information entropy measures for assembly line balancing optimization in the case of demand uncertainty.

2.1.2 Perceived assembly complexity

An assembly process may be perceived by humans as complex due to a wide range of variables (e.g., knowledge, personal experience, required capabilities, cognitive and physical effort required, etc.). In order to provide a comprehensive description of complexity many researchers used questionnaires and self-reporting tools. In this way, many different variables may be included in the assessment. Mattsson et al. (2014, 2016) proposed 5 main causes that may influence the operators' perception of assembly complexity (i.e., product variants, layout, work content, tools and information). A set of statements for each cause was provided and rated by workers on a 5-level scale. By aggregating these ratings, an overall complexity index (CXI) can be calculated. This index was further used in a practical case study of an automotive company (Mattsson *et al.* 2020). Similarly, Falck et al. (2017a) developed 16 basic complexity criteria, organized in five categories: knowledge demanding tasks, variety of fitting demands, many choice options, concentration/memory intensive tasks and physically/visually demanding tasks. Teams of experts determine the fulfilment of basic complexity criteria and provide an overall qualitative assessment on a 5-level complexity scale (Falck, Tarrar, *et al.* 2017, Falck, Örtengren, *et al.* 2017b).

2.2 Complexity in Human Robot Collaboration Assembly

Gervasi et al. (2020) underlined the importance to follow an holistic approach in evaluating HRC, taking into account various dimensions such as adaptivity, safety, human factors, team organization, knowledge, etc. Cobots, indeed, proved to be a support for humans both from an ergonomic and a cognitive point of view. Gualtieri et al. (2021) designed a novel collaborative assembly workstation that highly improved operators' physical ergonomics with respect to the traditional manual one. Regarding cognitive support, Buerkle et al. (2022) implemented a sensor framework for humans in HRC including both physiological, objective and subjective measures to assess perceived workload. Results showed that the robot generally reduced humans' perceived workload. Similarly, Gervasi et al. (Gervasi *et al.* 2023) compared manual assembly and collaborative assembly showing the support effect of cobots in reducing process failures, humans' perceived workload and stress.

Providing an overall and omni-comprehensive assessment of the perceived complexity in a manual assembly is a great challenge, as shown in section 2.1. A fortiori, in collaborative assemblies, such assessments become even more difficult as additional dimensions are involved, due to the interaction with the cobot. To the best of authors' knowledge, few studies investigated complexity assessment in HRC assembly processes:

- Malik and Bilberg (2019) who defined set of metrics grouped into 3 main categories (i.e., product, process and workspace) to assess the complexity of a collaborative assembly;
- Parsa and Saadat (2021) who developed an ordinal score-based methodology to assess the difficulty of performing disassembly tasks with cobots
- Capponi et al. (2022) who defined a theoretical framework of collaborative assembly complexity, considering product, operational and interaction complexity.
- Wang et al. (2022) who proposed an information entropy-inspired method to quantify complexity in collaborative assembly

All these studies recognised the importance of monitoring assembly complexity in enhancing manufacturing performances and operators' well-being. While some approaches investigated only specific aspects with a narrow focus on assembly complexity, others examined the interaction of multiple factors, such as product and process features, working environment characteristics, operator capabilities, ergonomics, etc. Despite these efforts, providing a holistic and quantitative assessment of perceived complexity remains a significant challenge.

3. Adapting Thurstone Law to assembly complexity assessment

The literature review section highlights that holistic methods to assess perceived assembly complexity are predominantly semi-quantitative and results are typically expressed on ordinal scales. With these scales, distances between objects are not defined (Stevens 1946). To overcome this limitation, a novel method to assess perceived assembly complexity, using the Thurstone law of comparative judgements (Thurstone 1927), is herein proposed.

3.1 Thurstone Law of Comparative judgements

The core of Thurstone's model is the concept of a "psychological continuum", which refers to an ideal space in which objects are placed on a one-dimensional scale based on their degree on a specific characteristic. The assessment of this attribute is qualitative and subjective, with different subjects providing their own judgements. The position of an object on the scale is directly related to the degree of the attribute it possesses, with the attribute increasing to the right and decreasing to the left of the scale (Franceschini and Maisano, 2020). One of the key contributions of Thurstone Law is the ability to create a scale with interval properties from data initially expressed as paired comparisons. In this work the Thurstone Law of comparative judgements known as “case V” will be addressed.

Consider a set of objects $O_1, \dots O_i, \dots O_n$ to be compared in pairs by a set of m experts. Thurstone law of comparative judgements – case V (Thurstone 1927, Franceschini and Maisano 2020) is based on the following assumptions:

- according to the concept of *modal discriminal process* (Thurstone, 1927), the position of an object in the continuum is described by a normal distribution ($O_i \sim N(\mu_i, \sigma_i^2), \forall i = 1, \dots, n$)
- all the objects have the same variance. i.e., $\sigma_1^2 = \sigma_i^2 = \sigma_n^2 = \sigma^2$
- equal Pearson correlation between all pairs of objects ($\rho_{ij} = \rho, \forall_{ij} = 1, \dots, n$ where ρ_{ij} is the Pearson correlation coefficient between object i and j)

Two possible outcomes may be obtained through pairwise comparisons: strict preference ($O_i > O_j$ or $O_i < O_j$), indifference ($O_i \sim O_j$). After all experts have compared all possible pairs of the n objects, a frequency matrix F can be computed. The element $f_{ij} \in F$ is defined as follows:

$$f_{ij} = |A| + 0.5|B| \tag{1}$$

where:

- “| |” is the cardinality operator that counts the number of elements in a given set,
- A is the sub-set of experts for which $O_i > O_j$,
- B is the sub-set of experts for which $O_i \sim O_j$.

From the F matrix, the proportion matrix P can be computed. The element $p_{ij} \in P$ represents the observed proportion in which $O_i > O_j$, defined as follows:

$$p_{ij} = \frac{f_{ij}}{m} \quad (2)$$

where:

- f_{ij} is obtained through Eq. (1)
- m represents the set of experts.

From the proportion matrix P, standard score z_{ij} between O_i and O_j can be computed by the following formula:

$$z_{ij} = \Phi^{-1}(1 - p_{ij}), \;\; z_{ij} \in Z \quad (3)$$

Where:

- p_{ij} is computed using Eq. (2)
- Φ represents the cumulative distribution function of the standard normal distribution.

It can be readily demonstrated that by summing the values in the j-*th* column of the Z matrix and dividing by n, the average position (μ_j) of the j-*th* object on the attribute continuum can be obtained (Franceschini and Maisano 2020). As shown in Figure 1, the obtained values are distributed accordingly to an interval scale (Stevens, 1946).

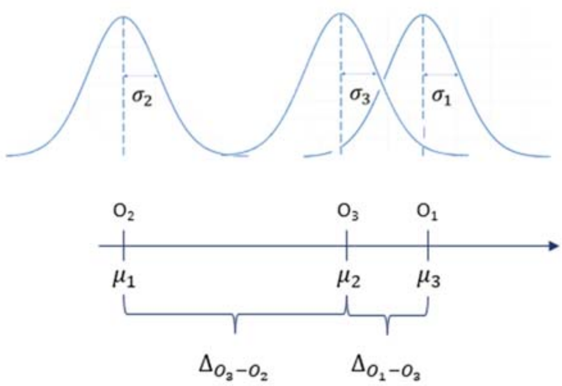


Figure 1 – Resulting objects' interval scale obtained through Thurstone Law of comparative judgements.

3.2. TICS method to assess perceived assembly complexity

Now let us suppose to assess the perceived complexity of n different assembly processes by m assembly operators (experts). The implementation of TICS method can be summarised in the following steps (as shown in Figure 2):

- Step 1: Assembly processes execution. Each operator performs all n different types of assembly processes.
- Step 2: Pairwise comparisons collection. Each operator is asked, by pairwise comparisons, to choose the assembly processes felt more complex to perform. For each operator, the number of required pairwise comparison is $C_{n,2} = \binom{n}{2}$.
- Step 3: Construction of the perceived complexity scale through the Thurstone Law of comparative judgements.

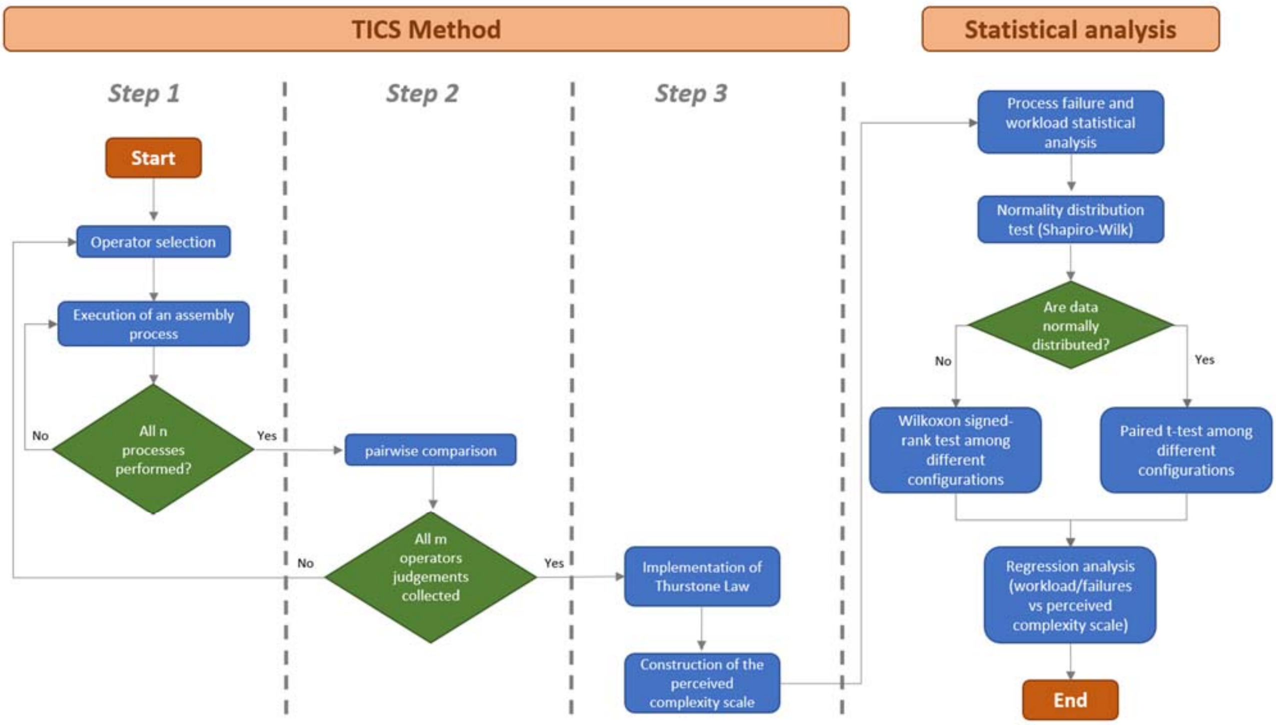


Figure 2 – Flowchart describing the steps of the proposed method.

The resulting perceived complexity scale can be related, for example, to the number of failures or to the operator's perceived workload. This aspect allows process designers to leverage perceived complexity in order to improve process and product quality and the operators' well-being (ElMaraghy *et al.* 2012, Genta *et al.* 2018, Verna *et al.* 2022b, 2022a). The following variables are considered in the analysis:

- Human-caused Process failures, i.e., the errors committed exclusively by humans during the assembly processes.
- Perceived workload, i.e., the subjective workload that operators experience in performing a task. The well-established tool "NASA-TLX" was considered in the study (Hart and Staveland 1988).

The following sections describe an experimental implementation of the proposed method. Statistical analysis will be deepened in section 5.2 where relations between the resulting perceived complexity scale and process quality variables (i.e., human-caused process failures and perceived workload) are then explored.

4. Case-study description

The case study concerns the manual and collaborative assembly of three different industrial products. The experiment involved 18 participants and took place in the "Mind4Lab" of Politecnico di Torino. For collaborative assemblies, the experiment was conducted using a collaborative Universal Robot UR3. All participants, aged between 20 and 25 years, were students of management and production engineering at Politecnico di Torino. None of the participants reported to have previous experience working with collaborative robots. Furthermore, none of them claimed to have performed a manual assembly process in laboratory or industrial settings. Participants with no prior experience were specifically selected as a deliberate methodological choice to control for prior knowledge or biases that could potentially influence the results of the experiment.

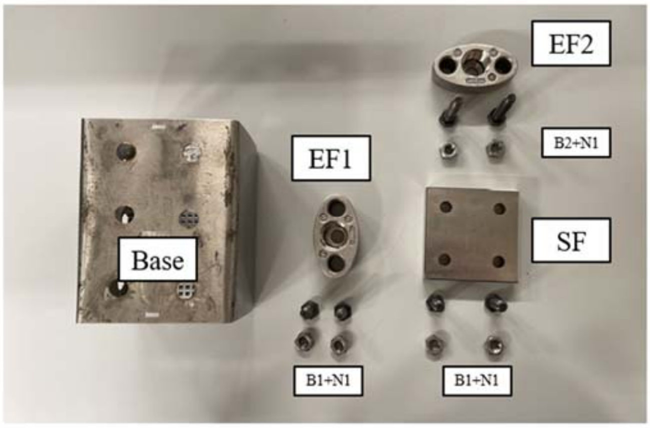
4.1 Experimental methodology

Participants were asked to complete an assembly process of three different products: a mechanical equipment ("1"), a tile cutter ("2") and a diaphragm water pump ("3") respectively in two modalities: manual (M) and collaborative (HRC). Hence, a total of 6 assembly process configurations were carried out, i.e., 1M, 1HRC, 2M, 2HRC, 3M and 3HRC. For each configuration, 2 training trials were planned and then, 4 repetitive trials were performed. Firstly, the decision of repeating an assembly process four times arose mainly from time constraints (this schedule involved each participant for four hours). Secondly, after a preliminary internal study, it was considered a reasonable trade-off to obtain an authentic assessment of perceived complexity. Too many trials could cause participants to become overly familiar with the task, thus distorting their evaluations toward a perception of simplicity. Conversely, too few trials could amplify the effect of initial unfamiliarity, leading to high perceived complexity.

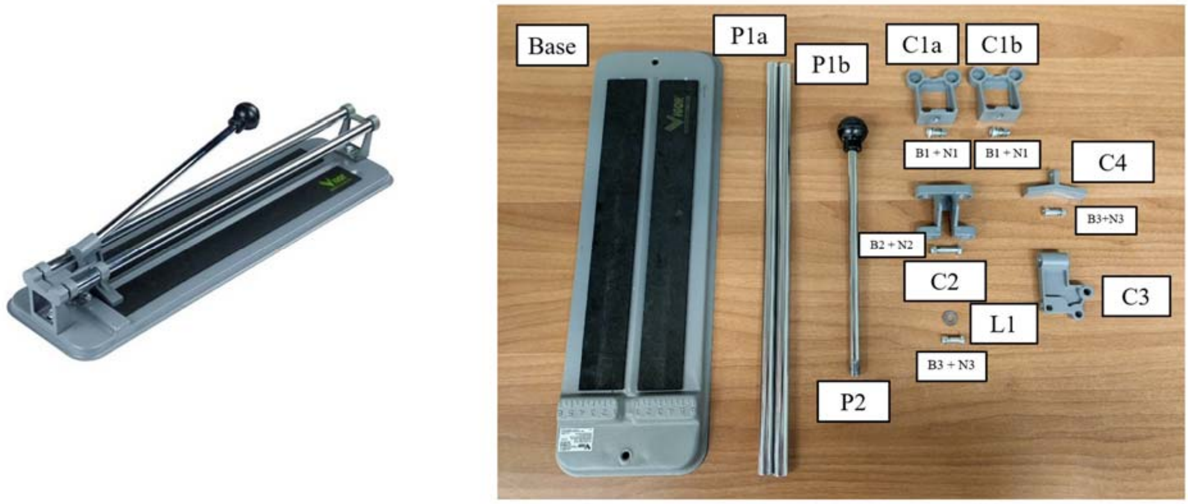
Before starting the experiment, participants were also provided with video assembly instructions illustrating step-by-step all the 6 tasks that needed to be performed. For each participant the experimental procedure was the following: firstly, after a brief introduction, the participant was shown the randomly selected configuration to perform. After two training trials, four repetitions for each configuration were carried out. Then, the participant was asked to fill out the NASA-TLX questionnaire. After all configurations were performed, participants were asked to pairwise compare the six configurations in terms of perceived complexity.

4.2 Products and assembly processes

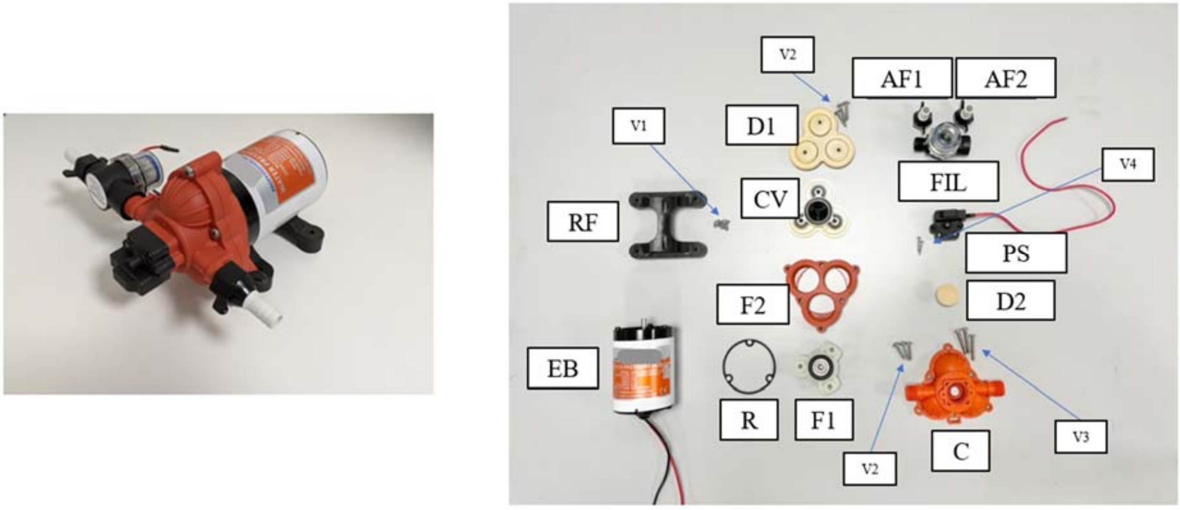
As anticipated, the case study considered three different products: (a) a simple mechanical equipment, (b) a tile cutter and (c) a diaphragm water pump (see Figure 3). The weight of the three assembled products is respectively: 1.45 kg, 1.23 kg and 1.8 kg while the three maximum dimensions are respectively: 125 x 141 x 96 mm, 440 x 90 x 140 mm and 265 x 123 x 203 mm. The code and description of each component are described in detail in Appendix A.



(a)



(b)



(c)

Figure 3 – The three products and their respective components: (a) mechanical equipment, (b) tile cutter and (c) diaphragm water pump

The list of the elementary tasks included in the assembly of the three reference products and their related allocation between agents (i.e., human, cobot) is provided in Appendix B. The cobot was programmed using the basic Move tool of the teach pendant provided by Universal Robot. Pick and place tasks of heavier and more rigid components were assigned to the cobot, as they could be too repetitive, and thus strenuous, for human operators. Fastening tasks, on the other hand, were assigned to the human operator, as they required greater flexibility. In manual modality all tasks were carried out by humans.

Figure 4 shows the assembly workstation and the three different assembly work-areas for the three reference products. The assembly work-area consisted of two sub-areas:

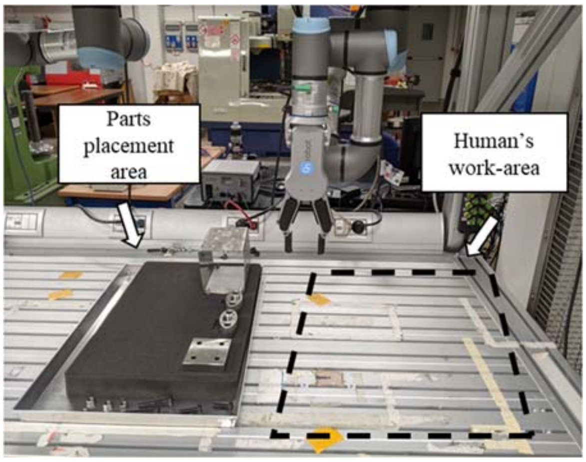
- the "parts placement area" where all components were arranged on a tray. From here, the cobot picked them and moved them within the "human's work-area".
- the "human's work-area" where the human operator performed manual tasks to complete the process.

The layout of each assembly work-area was different depending on the product to be assembled:

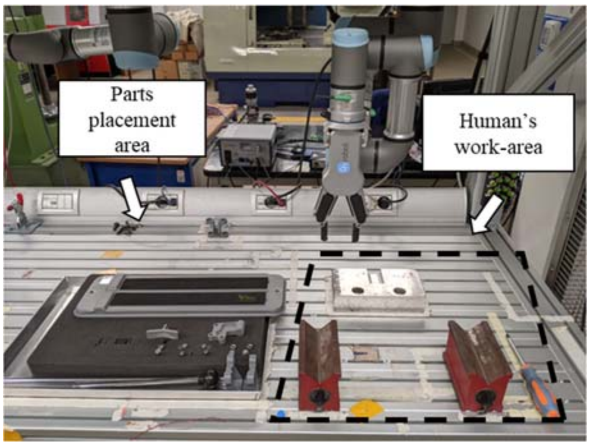
- mechanical equipment: all operations could be carried out without the aid of supports or screwdrivers (see Figure 4b)
- tile cutter: it was necessary to use physical supports to raise the base and facilitate both the screwing of the two bolts (i.e., B1) and the robot's grip. A screwdriver was also provided (see Figure 4c).
- diaphragm water pump: a screwdriver was provided, but no additional physical support was needed (see Figure 4d).

The parts to be assembled were arranged on a tray, in predefined positions. The operator stood in front of the assembly area and performed the assembly task, following the task allocation provided in appendix B.

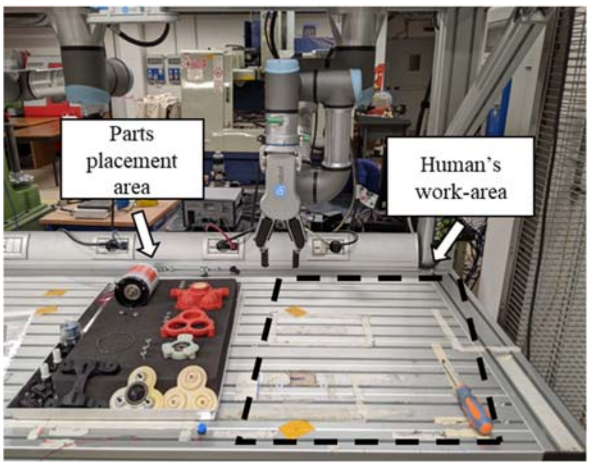
(a)



(b)



(c)



(d)

Figure 4 – The workstation (a) and the three assembly work-areas: (b) mechanical equipment, (c) tile cutter and (d) diaphragm water pump

4.3 Data collection

During the experiments, three main types of data were collected: pairwise comparisons, process failures and perceived workload by NASA-TLX tool. The process failure and NASA-TLX database consisted of 108 observations (i.e., 6 configurations for 18 participants). For each configuration, quantity and type of process failures and related NASA-TLX results were collected. For each participant the whole number of pairwise comparisons were $C_2^6 = 15$. Hence, a total of 270 (15x18) judgements were collected. Statistical analyses were performed using R©.

4.3.1 Pairwise comparisons

To implement Thurstone method the pairwise comparison on perceived complexity among the six configurations were collected. Participants were asked to choose the configuration they felt as more complex to perform. The exact question asked was: *"which of the two configurations was more complex for you to complete?"*.

4.3.2. Process failures

Process failures refer to possible errors caused by operators (both human and robot) during the assembly process. In general, process failures jeopardise the efficiency and productivity of an assembly process (Maisano *et al.* 2019, Gervasi *et al.* 2023). In this paper only human-caused process failures were considered to investigate the potential support of cobots. Based on the classification proposed by Gervasi et al. (2023), the following categories of human-caused process failures were considered:

- Wrong component/connectors selection: it refers to situations in which the operator picks up the wrong component or connector.
- Wrong component/connectors position: it occurs when an operator incorrectly positions a component.
- Incorrect assembly: it occurs when an operator incorrectly assembles a product.
- Dropping of components/connectors/tools: it refers to situations in which the operator drops components/connectors or tools.
- Part damage: it occurs when the operator damages a component or connector.
- Wrong input to cobot: it refers to cases in which the operator gives input to cobot when not necessary.

4.3.3. Perceived workload

To obtain data concerning the perceived workload experienced by participants, the NASA-TLX questionnaire (Hart and Staveland 1988) was adopted. NASA-TLX is a common tool used to rate perceived workload. It involves the assessment of six dimensions potentially influencing perceived workload (i.e., mental demand, physical demand, temporal demand, performance, effort and frustration) on 100-point scale (5-point steps). The six aforementioned dimension were then compared in pairs and participants were asked to choose the one more influencing their own perceived workload. By counting the number of times that each of the 6 dimensions is rated as more influential than one of the others, it is possible to calculate a weight for each dimension per participant, which is then used to obtain a weighted average workload (Hart and Staveland 1988).

5. Analysis of results

This section presents the main results of the experimental case-study organized as follows: subsection 5.1. describes how TICS method is implemented to obtain an overall perceived assembly complexity scale; subsection 5.2. investigates the statistical relationship among human-caused process failures, workload and the obtained perceived complexity scale.

5.1 Development of a complexity scale using Thurstone Law of comparative judgements

TICS method was applied to obtain a perceived complexity scale for the case study described in section 4. Figure 5 shows the average times distributions for each assembly process configuration. For each participant and configuration, the average assembly time over the 4 trials was computed. Hence, for the same configuration, each histogram in Figure 5 is based on 18 observations. As anticipated, various studies demonstrated the existence of a relationship between assembly time and both objective and perceived complexity (Alkan 2019, Sudhoff *et al.* 2022, Verna *et al.* 2023). The normality distribution of the 6 configurations was confirmed using Shapiro's test. This test has proved to be effective even for small sample size (Shapiro and Wilk 1965). The related p-values obtained were: $p_{1HRC} = 0.8389$; $p_{2HRC} = 0.7492$; $p_{3HRC} = 0.9905$; $p_{1M} = 0.9529$; $p_{2M} = 0.8353$ and $p_{3M} \cong 1$.

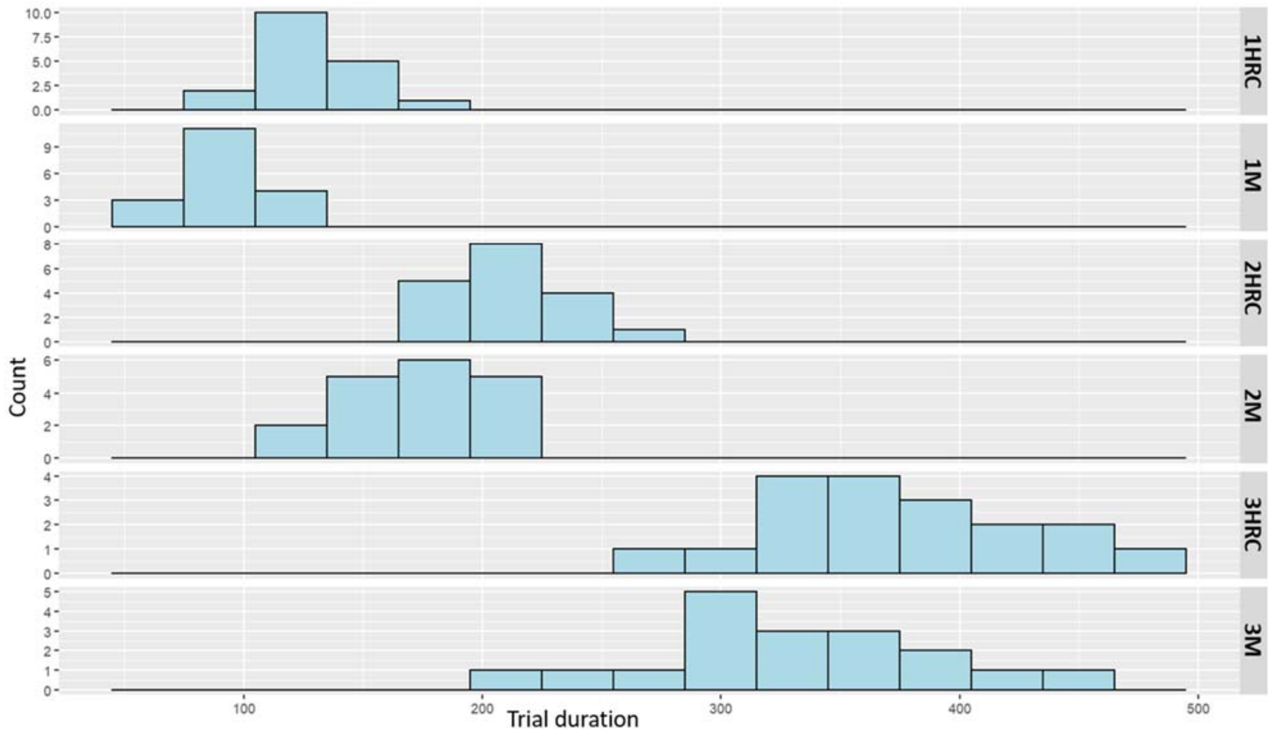


Figure 5 – Histogram plot of average assembly times for each configuration

Table 1 shows the values of mean, standard deviation and relative standard deviation for the average times (per participant) of the six different configurations. As could be expected, more complex products required higher assembly times and also presented greater variability, (Alkan 2019). Although standard deviation inevitably increased, the relative standard deviation among the six configurations remained similar. However, assembly times in collaborative modality are strongly influenced by how the task was scheduled and the cobot trajectories were defined. To simplify the problem, as a first approximation, it was decided not to analyse the relationships between assembly times and the perceived complexity scale. In subsequent experiments this aspect will be analysed in detail.

Table 1 – Average time, standard deviation and relative standard deviation of average assembly times for each configuration

Configuration	**Average time ($\hat{\mu}$)**	**Standard deviation ($\hat{\sigma}$)**	**Relative standard deviation ($\hat{\sigma}/\hat{\mu}$)**
1HRC	128.3 s	17.8 s	0.14
1M	94.4 s	15.7 s	0.17
2HRC	211.5 s	23.6 s	0.11
2M	172.9 s	30.7 s	0.18
3HRC	373.9 s	55.6 s	0.15
3M	330.0 s	60.2 s	0.18

5.1.1 TICS method: case-study implementation

The implementation of TICS method can be summarised in the following steps (see Section 3):

- Step 1: Each participants performed six different configurations (i.e., 1M, 1HRC, 2M, 2HRC, 3M, 3HRC).
- Step 2: Pairwise comparisons. After having performed all 6 configurations, each participant compared them in pairs and provided 15 judgements about perceived complexity, i.e., strict preference (">" or "< ") or indifference (" = ").
- Step 3: Construction of the perceived complexity scale using the Thurstone's Law of Comparative Judgement (see Table 2).
  1) Determination of the F matrix (see Table 2a) and P matrix (see Table 2b)
  2) Calculation of the Z matrix (see Table 2c). It should be noted that for values of p_{ij} equal to either 0 or 1, the resulting value of z_{ij} will be infinite (i.e., $z_{ij} = \pm\infty$). To overcame this issue, it was assumed that: $z_{ij} = \phi^{-1}(1-0.023) \approx 2$ if $p_{ij} \leq 0.023$; and $z_{ij} = \phi^{-1}(1-0.977) \approx -2$ if $p_{ij} \geq 0.977$ (Franceschini and Maisano 2020).
  3) Summing up values of Z matrix by column and dividing by n (number of configurations) the μ-values of each configuration were obtained (see Table 2c).

Table 2 – Results of Thurstone Law implementation in the experiment: (a) F matrix (where $f_{ij} = |A| + 0.5|B|$) regarding the paired comparison of each configuration, (b) P matrix (where $p_{ij} = \frac{f_{ij}}{m}$ and m is the number of participants) and (c) Z matrix (where $z_{ij} = \phi^{-1}(1-p_{ij})$) and final calculation of μ – values

(a)

F	1M	2M	3M	1HRC	2HRC	3HRC
1M	-	0	0	5.5	2	0
2M	18	-	5	16.5	15	5
3M	18	13	-	17	14	13.5
1HRC	12.5	1.5	1	-	2	0
2HRC	16	3	4	16	-	4.5
3HRC	18	13	4.5	18	13.5	-

(b)

P	1M	2M	3M	1HRC	2HRC	3HRC
1M	-	0.000	0.000	0.306	0.111	0.000
2M	1.000	-	0.278	0.917	0.833	0.278
3M	1.000	0.722	-	0.944	0.778	0.750
1HRC	0.694	0.083	0.056	-	0.111	0.000
2HRC	0.889	0.167	0.222	0.889	-	0.250
3HRC	1.000	0.722	0.250	1.000	0.750	-

(c)

Z	1M	2M	3M	1HRC	2HRC	3HRC
1M	0.000	1.995	1.995	0.508	1.221	1.995
2M	-1.995	0.000	0.589	-1.383	-0.967	0.589
3M	-1.995	-0.589	0.000	-1.593	-0.765	-0.674
1HRC	-0.508	1.383	1.593	0.000	1.221	1.995
2HRC	-1.221	0.967	0.765	-1.221	0.000	0.674
3HRC	-1.995	-0.589	0.674	-1.995	-0.674	0.000
$\boldsymbol{\mu}$	-1.286	-0.947	0.006	0.528	0.763	0.936
$\boldsymbol{\mu + D_{shift}}$	0	0.339	1.292	1.814	2.049	2.222
Δ	-	0.34	0.95	0.52	0.24	0.17

These dimensionless values define the perceived assembly complexity scale for the 6 different configurations considered (see Figure 6). The six different configurations are placed along a perceived complexity scale with interval properties (Stevens 1946). In this scale 1M represents the least complex perceived configuration, while 3M represents the most complex one.

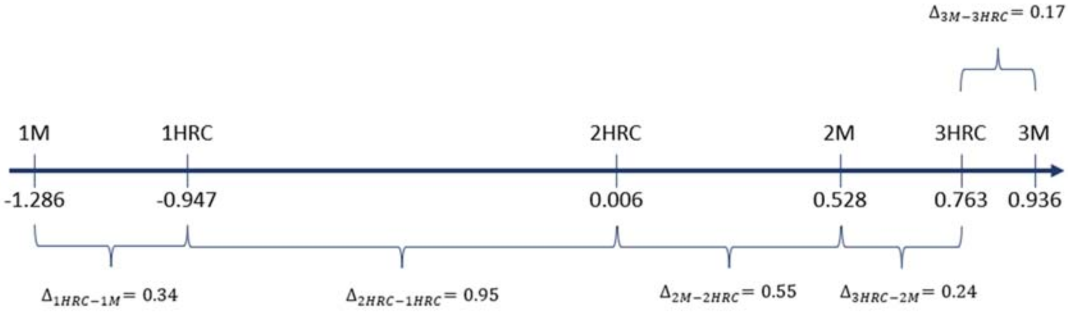


Figure 6 – Overall perceived assembly complexity scale obtained through TICS method.

To better analyse the results obtained, the lowest value of perceived complexity (i.e., 1M) was shifted to zero (i.e., $\boldsymbol{D_{shift}} = 1.286$) and all other average complexity values (i.e., μ) were shifted by the same amount to have only positive values. This operation complies with interval scale properties (Stevens 1946, Franceschini *et al.* 2019). Considering the unstructured feedback collected during the experimental campaign, three aspects can be remarked:

- In assemblies performed with the cobot (i.e., 2HRC and 3HRC) the complexity perceived by operators decreases if compared to the respective manual assembly. Indeed, the robot provided support to the operator by defining step by step the tasks to be performed. This is particularly valuable when assembly processes are very time consuming with many different operations to be carried out.
- In very simple product assemblies (e.g., mechanical equipment), robot support was not so relevant. The operator was able to perform few simple operations more efficiently and therefore perceived the robot as a useless support.

These preliminary considerations will, however, be discussed in more detail in the next subsections.

5.2 Statistical analysis

In this subsection the effects of assembly complexity on process failures and perceived workload are addressed. In detail, the aim of the authors was to investigate whether the perceived complexity scale represents a suitable proxy to describe the occurrence of human-caused process failures and perceived workload. Conceptually, one would expect that greater perceived complexity resulted in more process failures and a higher perceived workload. To this purpose, two different types of analysis were performed:

- Statistical hypothesis tests: Process failures and perceived workload of the six different configurations were compared using two statistical hypothesis test , i.e., paired t-tests, if the normality assumption wasn't rejected, or the Wilcoxon signed-rank test (Wilcoxon 1945), if the normality assumption was rejected. Wilcoxon signed-rank test is a non-parametric statistical test. It represents an alternative to paired t-test since it doesn't assume the normality distribution (Wilcoxon 1945). Given the small sample size, the normality assumption was tested using the Shapiro-Wilk test (Shapiro and Wilk 1965). Since each participant performed all six configurations, paired difference tests were performed in order to consider the within subject effect.

- Regression analysis: the aim of this analysis was to investigate whether the perceived complexity scale obtained through TICS method well described the occurrence of failures and the perceived workload.

5.2.1 Human-caused Process failures

As discussed above human-caused process failures are concerned with errors made by operators during the assembly processes. Number and types of failures were collected during the experiment. Two outlier observations were detected using the Inter-Quartile Range ("IQR") rule (Tuckey 1977) and thus excluded. For each configuration, a Shapiro-Wilk normality test was performed to test the distribution of the number of failures. The p-values obtained through the Shapiro-Wilk test were the following: $p_{1HRC} = 0.02818$; $p_{2HRC} = 0.04752$; $p_{3HRC} = 0.46$; $p_{1M} = 0.01029$; $p_{2M} = 0.7799$ and $p_{3M} = 0.006533$. In most configurations, the normality hypothesis was rejected, so a Wilcoxon signed-rank test was implemented. The matrix in Table 3 shows the adjusted p-values resulting from pairwise comparisons of the 6 configurations. Figure 7 shows the boxplots of Human-caused process failures for the 6 different configurations.

Data suggested that different products had a significant impact on process failures. Therefore, products with greater quantity and variety of parts and operations led to more process failures. All adjusted p-values related to the tile cutter and diaphragm water pump suggested significant difference in human process failures when compared with the assembly of the mechanical equipment, regardless of the presence of the cobot. For the same product, however, it was seen that the median of failures made in collaborative modality was slightly lower than in manual modality both for the tile cutter and the diaphragm water pump. Although not statistically significant, this result would suggest that the cobot supported humans in completing the assembly task, preventing potential failures. In the case of the mechanical component, instead, an opposite behaviour occurred, apparently suggesting the marginal role of the cobot in simple assembly processes where the human operator seemed to be more efficient.

Table 3 – Pairwise p-values of the Wilcoxon signed rank test for human-caused process failures ("" for p-values <0.05; "**" for p-values <0.01; "***" for p-values <0.001)*

p.adjust	***1HRC***	***1M***	***2HRC***	***2M***	***3HRC***	***3M***
1HRC	-					
1M	1,000	-				
2HRC	0.0067**	0.0219*	-			
2M	0.0030**	0.0030**	0.0883	-		
3HRC	0.0030**	0.0030**	0.0697	1.000	-	
3M	0.0030**	0.0030**	0.0503	1.000	1.000	-

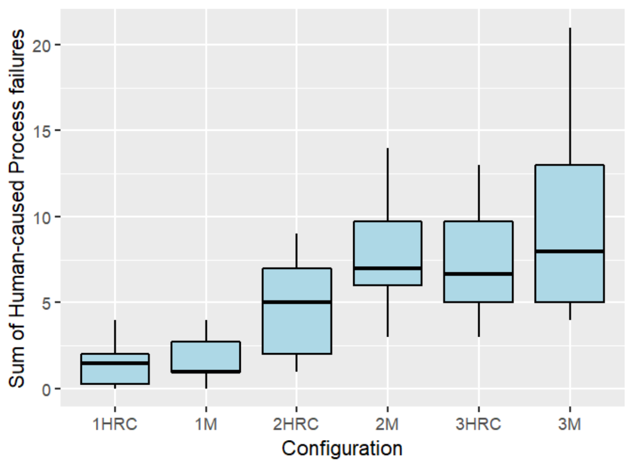
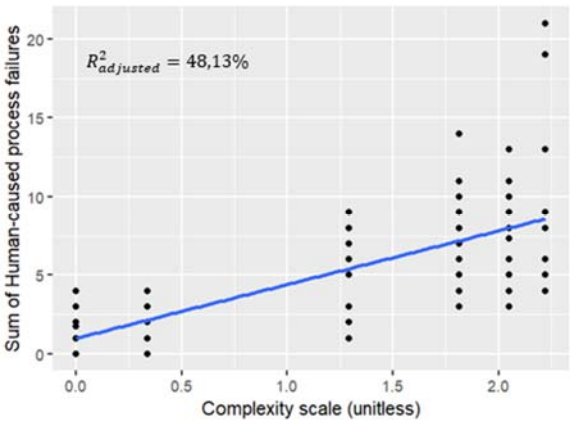


Figure 7 – Boxplot of human-caused process failures for the six considered configurations.

In order to analyse the relationship between perceived assembly complexity and the occurrence of process failures, a regression analysis was performed. The parameters obtained were:

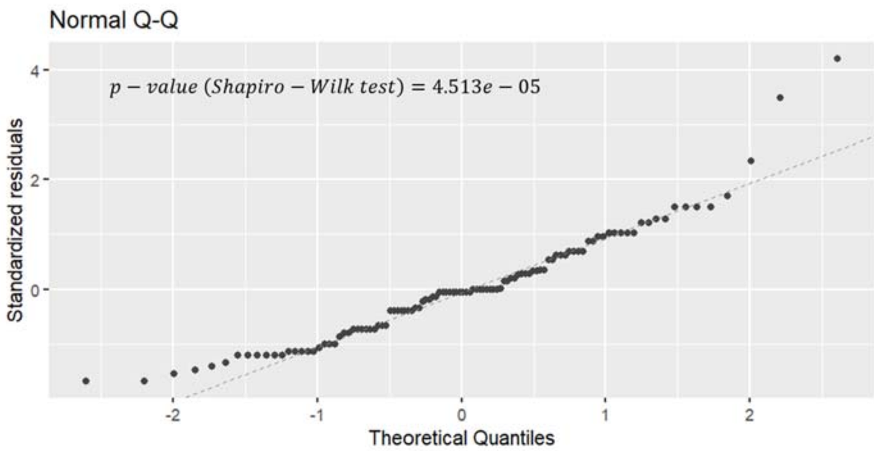
- Intercept: $a = 1.0072$ $(p - value = 0.0572)$
- Perceived complexity coefficient: $b = 3.4048$ $(p - value < 2e - 16)$.

Figure 8 shows the results of the linear regression analysis, and the related residuals normal Q-Q plot. Residuals distribution (see Figure 8b) was tested using the Shapiro-Wilk normality test that led to the rejection of the normality assumption $(p = 4.531e - 05)$. However, considering $\alpha = 0.05$ the coefficient of perceived complexity resulted statistically significant. This suggests the existence of a relationship between the complexity scale obtained and the occurrence of process failures. Furthermore, it can be seen that the mutual distances among the last three configurations were shorter if compared to the others, thus indicating similar perception of assembly complexity. This confirms also the fact that the relative boxplots of the human-caused process failures overlapped in terms of median and variability and that the mutual differences did not result statistically significant.



Configuration	μ
1M	0
1HRC	0.339
2HRC	1.292
2M	1,814
3HRC	2.049
3M	2.222

(a)



(b)

Figure 8 – Regression analysis results: (a) regression plot of human-caused process failures vs perceived assembly complexity scale (TICS method) and (b) related normal Q-Q plot of residuals.

5.2.2 Perceived workload

Perceived assembly complexity also impacts on the workload required to perform an assembly process. To measure perceived workload NASA-TLX (Hart and Staveland 1988) was implemented. The analysis took into account the overall workload value obtained as a weighted average of six different dimensions (see section 4.3.2.). Similar to previous sub-section, two outliers were identified using 1,5IQR method and a Shapiro-Wilk normality test was performed. The p-values obtained through the Shapiro-Wilk test were the following: $p_{1HRC} = 0.1169$; $p_{2HRC} = 0.2235$; $p_{3HRC} = 0.3159$; $p_{1M} = 0.3098$; $p_{2M} = 0.2763$ and $p_{3M} = 0.3262$. The normality hypothesis for perceived workload data could not be rejected. Hence, paired t-test were implemented to analyse statistical differences among the 6 configurations. The related results are provided in Table 4 while Figure 9 shows the respective boxplots.

Table 4 – Pairwise p-values of the t-test ("" for p-values <0.05; "**" for p-values <0.01; "***" for p-values <0.001)*

p.adjust	*1HRC*	*1M*	*2HRC*	*2M*	*3HRC*	*3M*
1HRC	-					
1M	0.19531	-				
2HRC	0.08902	0.00239**	-			
2M	0.00149**	0.00034***	0.19531	-		
3HRC	4.8e-06***	8.7e-08***	0.00021***	0.04113*	-	
3M	3.0e-05***	9.9e-08***	0.00179**	0.01799*	0.19531	-

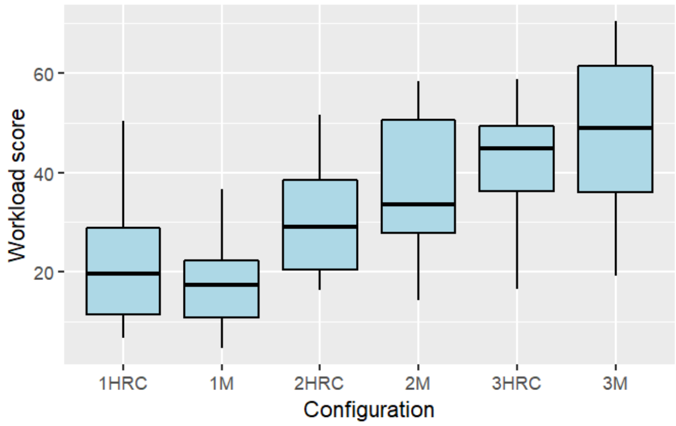


Figure 9 – Boxplot of perceived workload for the six configurations

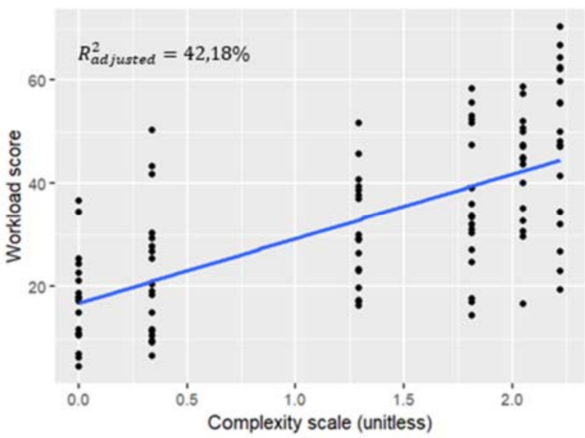
As in the case of process failures, the effect of the assembled product was significant for workload. Varying the assembly product, the participants observed different workloads. Specifically, products with more components required higher exertion and concentration from the operator. Regarding the support provided by the cobot in assembly processes, again, for very simple products (mechanical equipment), the presence of the robot led to higher workloads, especially in terms of perceived frustration. From the unstructured feedback collected, in fact, during the assembly of the mechanical equipment the cobot was perceived as useless, since it slowed down tasks that humans would have completed more quickly and efficiently. For more complex products, on the other hand, the cobot actually supported the human operator. By timing the various steps of the assembly process and providing the right component to be used, it allowed the operator to perform the correct assembly sequences. This resulted in fewer errors and, at the same time, less cognitive and physical effort.

The relationship between perceived assembly complexity and perceived workload was investigated performing a linear regression (see Figure 10). The obtained parameters of the regression analysis were:

- Intercept : $a = 16.765$ ($p - value = 5.91e - 12$)
- Perceived complexity coefficient: $b = 12.499$ ($p - value = 1.75e - 14$).

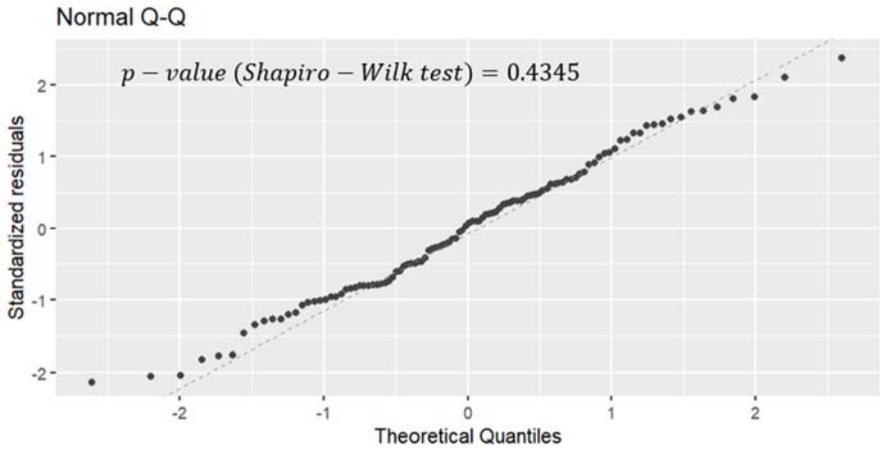
and resulted both statistically significant ($\alpha = 0.05$).

Unlike human-caused process failures, in addition to the significance of the parameters, the residuals of the linear regression presented a normal distribution (see Figure 10b). It can be deduced that a linear regression well approximated the evolution of perceived workload. Furthermore, these results emphasised that perceived workload could be used as a measure of the perceived complexity of operators. On the other hand, the value of $R^2_{adjusted}$ was found to be low, but this can be explained by the fact that data collected via subjective questionnaires generally exhibited high variability (see Figure 9).



Configuration	μ
1M	0
1HRC	0.339
2HRC	1.292
2M	1,814
3HRC	2.049
3M	2.222

(a)



(b)

Figure 10 – Regression analysis results: (a) regression plot of perceived workload vs perceived assembly complexity scale (TICS method) and (b) related normal Q-Q plot of residuals.

5.2.3 Comparison between objective and perceived complexity models

This section briefly shows a comparison between the perceived complexity scale obtained through the TICS method with the objective complexity model proposed by Samy and ElMaraghy H. (Samy and ElMaraghy 2010). Samy's method relates the assembly complexity of a product to the variety and quantity of its components and connectors and their geometric characteristics. Samy and ElMaraghy therefore proposed a product assembly complexity index (i.e., $C_{product}$) defined as follows:

$$C_{product} = \left[\frac{n_p}{N_p} + CI_{product}\right][\log_2(N_p + 1)] + \left[\frac{n_s}{N_s}\right][\log_2(N_s + 1) \qquad (4)$$

Where:

- n_p is the number of unique parts and N_p is the total number of parts composing a product.
  n_s is the number of unique fasteners and N_s is the total number of fasteners. In this work a single bolt and its related nut was considered as a single fastener.
- $CI_{product} \in [0; 1]$ is a complexity index related to geometrical and dimensional features of parts. It can be computed using manual handling and joining difficulty factors derived from Design for Assembly (Samy and ElMaraghy 2010).

For simplicity, collaborative assemblies were neglected in this comparison. Therefore, three manual assembly configurations (i.e., 1M, 2M and 3M) were assessed using TICS method. To each configuration corresponds a specific objective complexity value ($C_{product}$) for manual assembly. Table 5 shows the main results obtained for the three analysed products.

Table 5 – Results comparison between TICS and Samy and ElMaraghy H.'s method (Δ_{i-1M} represents numerical distances of the three configurations with respect to 1M value)

	Perceived complexity		**Objective complexity**						
	TICS method		***Samy and ElMaraghy H.***						
Configuration	$\boldsymbol{\mu}$	Δ_{i-1M}	N_p	n_p	N_s	n_s	$CI_{product}$	$C_{product}$	Δ_{i-1M}
1 M	-1.33	-	4	3	6	2	0.668	4.22	-
2 M	0.47	1.80	10	8	5	3	0.682	6.67	2.45
3 M	0.86	2.19	13	12	13	4	0.693	7.33	3.11

The method by Samy and ElMaraghy H. and TICS method led to very similar results. In fact, the ordering between the 3 assembly processes remained unchanged:

$$C_{mechanical\ equipment} < C_{tile\ cutter} < C_{diaphragm\ water\ pump}$$

$$\mu_{1M} < \mu_{2M} < \mu_{3M}$$

In terms of both objective complexity and perceived complexity, the greatest distance could be noted between the mechanical equipment and the tile cutter, while it decreased between the tile cutter and the pump. Figure 11 shows the results of TICS method and the corresponding complexity value using Samy and ElMaraghy's method. A preliminary linear trend can be observed between values of the two reference methods.

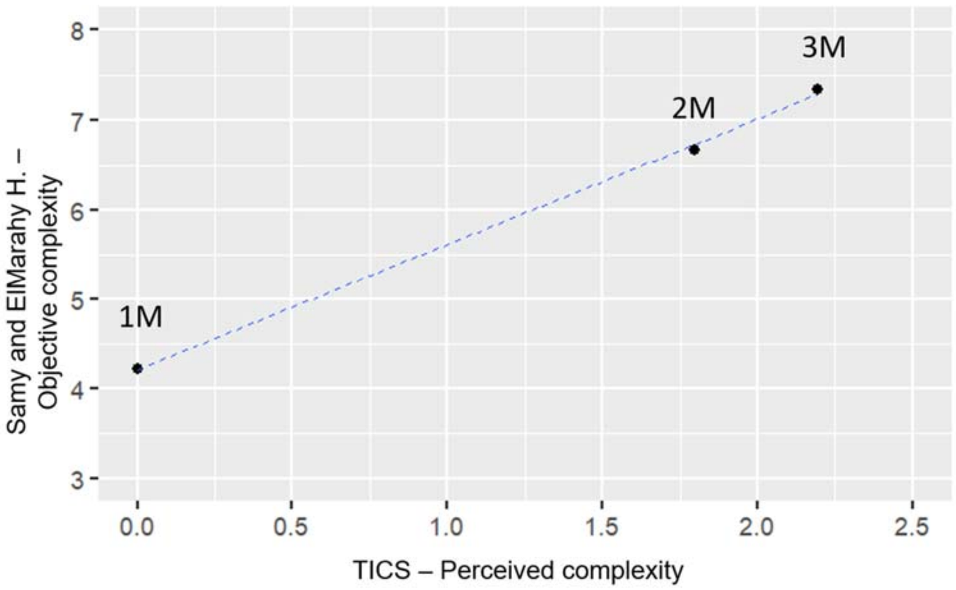


Figure 11 – Samy and ElMaraghy H.'s objective complexity vs TICS perceived complexity

Hence, albeit for this small sample, results showed a general concordance between perceived and objective complexity. Obviously, such comparisons should be generalized, considering more products.

6. Conclusions

This paper proposes a novel method (named TICS) to define an assembly perceived complexity scale for both manual and collaborative processes. The method is based on the application of the Thurstone Law of Comparative judgement. The main contribution of this methods is the possibility to create a link between perceived complexity and respectively process failures and perceived workloads. To this end, an experimental case-study concerning both manual and collaborative assemblies of three different products was proposed. The results showed that the Thurstone-inspired method is related both to the perceived workload of operators and process failures.

A second objective was to verify the potential impact of the cobot's presence in collaborative assemblies with respect to manual processes. Results showed that in terms of both failures and workload, the cobot only supported the operator in more complex assembly processes. For very simple products, such support was not noticeable.

The implementation of such methodology can lead to useful benefits in industrial context. They can be subdivided into three categories: process improvements, workers' training and cobot integration decision-making:

- Process improvements: unlike many methods in the literature, TICS allows the creation of a quantitative scale of perceived assembly complexity, which can be correlated with typical process control parameters such as product defects and process errors. Furthermore, this methodology can also be adopted to identify those assembly processes in which actions should be taken to enhance humans' well-being.
- Training: process designers can use this tool to prioritise tasks perceived as more complex, and therefore those that need more emphasis in training so that workers can be adequately prepared.
- Cobot integration decision-making: in this specific case, TICS makes it possible to identify and quantify the support of a cobot. It can therefore be used as a preliminary decision-making tool to shed light on which processes are worth investing in collaborative robotics.

One limitation of the proposed methodology is that it is based on "a posteriori" assessments, i.e., it can only be implemented after an operator has already performed the assembly process. Furthermore, this study was conducted in a laboratory setting, which may only partially replicate a real industrial context. Secondly, more products, and thus assembly processes, should be tested in order to have more robust and generalizable results. Finally, the current paucity of literature regarding complexity on collaborative assembly makes it difficult to find similar methodologies for a meaningful comparison with the one proposed in this paper.

Future developments will concern:

- the extension of the experiment to experienced participants and an increase in the number of repetitive trials to investigate whether complexity assessments are influenced by prior expertise and learning effects;
- the definition of an "a priori" assembly complexity model, specifically for collaborative assemblies, which will allow prediction of process failures and perceived workload
- the analysis of the efficiency gains from the use of collaborative robots in manufacturing processes, measuring both improvements in production speed, reduction in errors, overall workflow efficiency, costs and their contextual impact on perceived complexity.

Finally, another area worth investigating concerns how new technologies (augmented and virtual reality) might affect perceived assembly complexity in industrial processes.

References

Alkan, B., 2019. An experimental investigation on the relationship between perceived assembly complexity and product design complexity. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 13 (3), 1145–1157.

Alkan, B., Vera, D.A., Ahmad, M., Ahmad, B., and Harrison, R., 2018. Complexity in manufacturing systems and its measures: a literature review. *European Journal of Industrial Engineering*, 12 (1), 116–150.

Ameri, F., Summers, J., Mocko, G., and Porter, M., 2008. Engineering design complexity: An investigation of methods and measures. *Research in Engineering Design*, 19, 161–179.

Battaïa, O., Dolgui, A., Heragu, S.S., Meerkov, S.M., and Tiwari, M.K., 2018. Design for manufacturing and assembly/disassembly: joint design of products and production systems. *International Journal of Production Research*, 56 (24), 7181–7189.

Bauer, A., Wollherr, D., and Buss, M., 2008. Human–robot collaboration: a survey. *International Journal of Humanoid Robotics*, 05 (01), 47–66.

Boothroyd, G., 1994. Product design for manufacture and assembly. *Computer-Aided Design*, 26 (7), 505–520.

Boothroyd, G. and Alting, L., 1992. Design for Assembly and Disassembly. *CIRP Annals*, 41 (2), 625–636.

Buerkle, A., Matharu, H., Al-Yacoub, A., Lohse, N., Bamber, T., and Ferreira, P., 2022. An adaptive human sensor framework for human–robot collaboration. *The International Journal of Advanced Manufacturing Technology*, 119 (1), 1233–1248.

Capponi, M., Mastrogiacomo, L., Antonelli, D., and Franceschini, F., 2022. Product complexity and quality in assembly processes: state-of-the-art and challenges for Human-Robot Collaboration. *In*: *Proceedings book of 5th International Conference on Quality Engineering and Management*. University of Minho, Portugal, 142–167.

Capponi, M., Mastrogiacomo, L., and Franceschini, F., 2023. General remarks on the entropy-inspired MCAT (Manufacturing Complexity Assessment Tool) model to assess product assembly complexity. *article in press*.

ElMaraghy, W., ElMaraghy, H., Tomiyama, T., and Monostori, L., 2012. Complexity in engineering design and manufacturing. *CIRP Annals*, 61 (2), 793–814.

ElMaraghy, W.H. and Urbanic, R.J., 2003. Modelling of Manufacturing Systems Complexity. *CIRP Annals*, 52 (1), 363–366.

ElMaraghy, W.H. and Urbanic, R.J., 2004. Assessment of Manufacturing Operational Complexity. *CIRP Annals*, 53 (1), 401–406.

Eskilander, S., 2001. Design for automatic assembly - a method for product design: DFA2. Ph.D. dissertation, Royal Institute of Technology, Sweden.

Falck, A.-C., Örtengren, R., Rosenqvist, M., and Söderberg, R., 2017a. Basic complexity criteria and their impact on manual assembly quality in actual production. *International Journal of Industrial Ergonomics*, 58, 117–128.

Falck, A.-C., Örtengren, R., Rosenqvist, M., and Söderberg, R., 2017b. Proactive assessment of basic complexity in manual assembly: development of a tool to predict and control operator-induced quality errors. *International Journal of Production Research*, 55 (15), 4248–4260.

Falck, A.-C., Tarrar, M., Mattsson, S., Andersson, L., Rosenqvist, M., and Söderberg, R., 2017. Assessment of manual assembly complexity: a theoretical and empirical comparison of two methods. *International Journal of Production Research*, 55 (24), 7237–7250.

Franceschini, F., Galetto, M., and Maisano, D., 2019. *Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators*. Springer.

Franceschini, F. and Maisano, D., 2020. Adapting Thurstone's Law of Comparative Judgment to fuse preference orderings in manufacturing applications. *Journal of Intelligent Manufacturing*, 31 (2), 387–402.

Fujimoto, H., Ahmed, A., Iida, Y., and Hanai, M., 2003. Assembly Process Design for Managing Manufacturing Complexities Because of Product Varieties. *International Journal of Flexible Manufacturing Systems*, 15 (4), 283–307.

Genta, G., Galetto, M., and Franceschini, F., 2018. Product complexity and design of inspection strategies for assembly manufacturing processes. *International Journal of Production Research*, 56 (11), 4056–4066.

Gervasi, R., Capponi, M., Mastrogiacomo, L., and Franceschini, F., 2023. Manual assembly and Human–Robot Collaboration in repetitive assembly processes: a structured comparison based on human-centered performances. *The International Journal of Advanced Manufacturing Technology*, 126 (3), 1213–1231.

Gervasi, R., Mastrogiacomo, L., and Franceschini, F., 2020. A conceptual framework to evaluate human-robot collaboration. *The International Journal of Advanced Manufacturing Technology*, 108 (3), 841–865.

Gualtieri, L., Rauch, E., and Vidoni, R., 2021. Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, 67, 101998.

Hart, S.G. and Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *In*: P.A. Hancock and N. Meshkati, eds. *Advances in Psychology*. North-Holland, 139–183.

Hinckley, C.M., 1994. A global conformance quality model. A new strategic tool for minimizing defects caused by variation, error, and complexity. Ph.D. dissertation, Stanford University, United States.

Hoffman, G., 2019. Evaluating Fluency in Human–Robot Collaboration. *IEEE Transactions on Human-Machine Systems*, 49 (3), 209–218.

Hvam, L., Hansen, C.L., Forza, C., Mortensen, N.H., and Haug, A., 2020. The reduction of product and process complexity based on the quantification of product complexity costs. *International Journal of Production Research*, 58 (2), 350–366.

Liu, X., Yang, X., and Lei, M., 2021. Optimisation of mixed-model assembly line balancing problem under uncertain demand. *Journal of Manufacturing Systems*, 59, 214–227.

Madappilly, P.J. and Mork, O.J., 2021. Review and modification of DFA2 methodology to support design for automatic assembly (DFAA) in the maritime industry. *Procedia CIRP*, 100, 744–749.

Maddikunta, P.K.R., Pham, Q.-V., B, P., Deepa, N., Dev, K., Gadekallu, T.R., Ruby, R., and Liyanage, M., 2022. Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, 26, 100257.

Maisano, D.A., Antonelli, D., and Franceschini, F., 2019. Assessment of Failures in Collaborative Human-Robot Assembly Workcells. *In*: *20th Working Conference on Virtual Enterprises (PRO-VE)*. Turin, Italy, 562–571.

Malik, A.A. and Bilberg, A., 2019. Complexity-based task allocation in human-robot collaborative assembly. *Industrial Robot: the international journal of robotics research and application*, 46 (4), 471–480.

Mattsson, S., Karlsson, M., Gullander, P., Van Landeghem, H., Zeltzer, L., Limère, V., Aghezzaf, E.-H., Fasth, Å., and Stahre, J., 2014. Comparing quantifiable methods to measure complexity in assembly. *International Journal of Manufacturing Research*, 9 (1), 112–130.

Mattsson, S., Tarrar, M., and Fast-Berglund, Å., 2016. Perceived production complexity – understanding more than parts of a system. *International Journal of Production Research*, 54 (20), 6008–6016.

Mattsson, S., Tarrar, M., and Harari, N., 2020. Using the compleXity index for improvement work: investigating utilisation in an automotive company. *International Journal of Manufacturing Research*, 15, 3.

Parsa, S. and Saadat, M., 2021. Human-robot collaboration disassembly planning for end-of-life product disassembly process. *Robotics and Computer-Integrated Manufacturing*, 71, 102170.

Roulet-Dubonnet, O., Sandøy, R.K., and Schulte, K.Ø., 2018. Case study: Application of Design for Automated Assembly methods in the development of an electronic product from early design to design freeze. *Procedia CIRP*, 70, 192–197.

Samy, S.N. and ElMaraghy, H., 2010. A model for measuring products assembly complexity. *International Journal of Computer Integrated Manufacturing*, 23 (11), 1015–1027.

Samy, S.N. and ElMaraghy, H., 2012. A model for measuring complexity of automated and hybrid assembly systems. *The International Journal of Advanced Manufacturing Technology*, 62 (5), 813–833.

Shannon, C.E., 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27 (4), 623–656.

Shapiro, S.S. and Wilk, M.B., 1965. An analysis of variance test for normality (complete samples)†. *Biometrika*, 52 (3–4), 591–611.

Shibata, H., 2002. Global assembly quality methodology: A new method for evaluating assembly complexities in globally distributed manufacturing. Ph.D. dissertation, Stanford University, United States.

Sigurjónsson, V., Johansen, K., and Rösiö, C., 2022. Exploring the operator's perspective within changeable and automated manufacturing – A literature review. *Procedia CIRP*, 107, 369–374.

Sinha, K., 2014. Structural complexity and its implications for design of cyber-physical systems. Ph.D. dissertation, Massachusetts Institute of Technology, United States.

Sinha, K. and de Weck, O.L., 2014. Structural Complexity Quantification for Engineered Complex Systems and Implications on System Architecture and Design. *In*: *Proceedings of the ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Portland, Oregon, USA.

Stevens, S.S., 1946. On the Theory of Scales of Measurement. *Science (New York, N.Y.)*, 103 (2684), 677–680.

Su, Q., Liu, L., and Whitney, D.E., 2010. A Systematic Study of the Prediction Model for Operator-Induced Assembly Defects Based on Assembly Complexity Factors. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 40 (1), 107–120.

Sudhoff, M., Schüler, P., Herzog, M., and Kuhlenkötter, B., 2022. Proving the Applicability of Assembly Complexity Measures for Process Time Prediction of Customer-specific Production. *Procedia CIRP*, 107, 381–386.

Sun, H. and Fan, S., 2018. Car sequencing for mixed-model assembly lines with consideration of changeover complexity. *Journal of Manufacturing Systems*, 46, 93–102.

Thurstone, L.L., 1927. A law of comparative judgment. *Psychological Review*, 34, 273–286.

Tuckey, J., 1977. *Exploratory data analysis*. Addison-Wesley.

Verna, E., Genta, G., and Galetto, M., 2023. A new approach for evaluating experienced assembly complexity based on Multi Expert-Multi Criteria Decision Making method. *Research in Engineering Design*, 34 (3), 301–325.

Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022a. Defect prediction for assembled products: a novel model based on the structural complexity paradigm. *The International Journal of Advanced Manufacturing Technology*, 120 (5), 3405–3426.

Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022b. Defects-per-unit control chart for assembled products based on defect prediction models. *The International Journal of Advanced Manufacturing Technology*, 119 (5–6), 2835–2846.

Villani, V., Pini, F., Leali, F., and Secchi, C., 2018. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55, 248–266.

Wang, H. and Hu, S.J., 2010. Manufacturing complexity in assembly systems with hybrid configurations and its impact on throughput. *CIRP Annals*, 59 (1), 53–56.

Wang, H., Wang, H., and Hu, S.J., 2013. Utilizing variant differentiation to mitigate manufacturing complexity in mixed-model assembly systems. *Journal of Manufacturing Systems*, 4 (32), 731–740.

Wang, Y., Wang, J., Feng, J., Liu, J., and Liu, X., 2022. Integrated task sequence planning and assignment for human–robot collaborative assembly station. *Flexible Services and Manufacturing Journal*.

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1 (6), 80–83.

Zanchettin, A.M., Ceriani, N.M., Rocco, P., Ding, H., and Matthias, B., 2016. Safety in human-robot collaborative manufacturing environments: Metrics and control. *IEEE Transactions on Automation Science and Engineering*, 13 (2), 882–893.

Zeltzer, L., Aghezzaf, E.-H., and Limère, V., 2017. Workload balancing and manufacturing complexity levelling in mixed-model assembly lines. *International Journal of Production Research*, 55 (10), 2829–2844.

Zhu, X., Hu, S.J., Koren, Y., and Marin, S.P., 2008. Modeling of Manufacturing Complexity in Mixed-Model Assembly Lines. *Journal of Manufacturing Science and Engineering*, 130 (5).

DECLARATIONS

Funding

This study was carried out within the MICS (Made in Italy – Circular and Sustainable) Extended Partnership and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.3 – D.D. 1551.11-10-2022, PE00000004). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Competing interests

The authors declare that they have no conflict of interest.

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by M. Capponi and R. Gervasi. The first draft of the manuscript was written by M. Capponi and R. Gervasi under the supervision of L. Mastrogiacomo and F. Franceschini. All authors read and approved the final manuscript.

Compliance with Ethical Standards

The authors respect the Ethical Guidelines of the Journal. Informed consent was obtained from all individual participants included in the study.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Appendix A – List of the components of the three reference products

Product	Parts and fasteners	Code	Quantities
Mechanical equipment	Base	Base	1
	Elliptical flange	EF1/EF2	2
	Square flange	SF	1
	Bolt type 1	B1	4
	Bolt type 2	B2	2
	Nuts type 1	N1	6
Tile cutter	Base	Base	1
	Lateral support	C1a/C1b	2
	Joint component	C2	1
	Cutting component	C3	1
	Blade	L1	1
	Tile blocker	C4	1
	Rail rod	P1a/P1b	2
	Handle	P2	1
	Bolt type 1	B1	2
	Bolt type 2	B2	1
	Bolt type 3	B3	2
	Nuts type 1	N1	2
	Nuts type 2	N2	1
	Nuts type 3	N3	2
Diaphragm water pump	Engine block	EB	1
	Rubber feet	RF	1
	Ring	R	1
	Flange 1	F1	1
	Flange 2	F2	1
	Diaphragm	D1	1
	Cover with valves	CV	1
	Cover	C	1
	Pressure switch	PS	1
	Pressure switch diaphragm	D2	1
	Filter	FIL	1
	Flow adapter	AF1/AF2	2
	Screws type 1	V1	2
	Screws type 2	V2	6
	Screws type 3	V3	3
	Screws type 4	V4	2

Appendix B – List and allocation of the elementary tasks concerning the assembly of the three reference products

Product	ID	Elementary task	Collaborative assembly process		Manual assembly process
			Human	Cobot	Human
Mechanical equipment	1	Pick and place BASE		X	X
	2	Pick and place EF1		X	X
	3	Screwing EF1 with Base	X		X
	4	Pick and place SF		X	X
	5	Screwing SF with Base	X		X
	6	Pick and place EF2		X	X
	7	Screwing EF2 with Base	X		X
	8	Pick the final product and place out of the assembly area		X	X
Tile cutter	1	Pick and place Base		X	X
	2	Pick and place C1a and C1b on Base	X		X
	3	Preliminary screwing C1a and C1b on Base	X		X
	4	Placing the subassembly (Base+C1a+C1b) out of the assembly area		X	X
	5	Pick and place C2		X	X
	6	Pick and place C3 in C2	X		X
	7	Screwing C3 and C2	X		X
	8	Pick and place L1	X		X
	9	Screwing L1 and C3	X		X
	10	Pick and place C4 in C3	X		X
	11	Screwing C4 and C3	X		X
	12	Placing the subassembly (C2+C3+C4+L1) out of the assembly area		X	X
	13	Pick and place subassembly (Base+C1a+C1b) back in the assembly area		X	X
	12	Insert sub-assembly (C2+C3+C4+L1) in both P1a/P1b	X		X
	13	Insert P1a/P1b in C1a/C1b	X		X
	14	Final screwing C1a/C1b on Base	X		X
	15	Pick and place P2	X		X
	16	Screwing P2	X		X
	17	Pick the final product and place out of the assembly area		X	X
Diaphragm water pump	1	Pick and place RF	X		X
	2	Pick and place EB		X	X
	3	Screwing EB with RF	X		X
	4	Pick and place F1		X	X
	5	Pick and place F2		X	X
	6	Insert F1 in F2	X		X
	7	Pick and place D1 on sub-assembly F1+F2	X		X
	8	Screwing D1, F1 and insert CV on D1	X		X
	9	Pick and place C		X	X
	10	Screwing C and F2	X		X
	11	Insert R on EB	X		X
	12	Insert and screwing sub-assembly pump head on EB (joining F1-EB)	X		X
	13	Pick and place D2 and PS on C	X		X
	14	Screwing PS and C	X		X
	15	Pick and place FIL	X		X
	16	Screwing FIL	X		X
	17	Pick and place AF1 and AF2	X		X
	18	Screwing AF1 and AF2	X		X
	19	Pick the final product and place out of the assembly area		X	X

List of Figures (Alt. text)

Figure 1 Alt text: Positions of three objects along an axis according to their mean and, for each of them, also the relative normal distribution is shown.

Figure 2 Alt text: Block diagram explaining all the activities of the proposed TICS method grouped by three main steps: assembly execution, pairwise comparison collection and construction of the Thurstone-inspired scale.

Figure 3 Alt text: Three photos of the assembled products: (a) mechanical equipment, (b) tile cutter and (c) diaphragm water pump, and their components marked with the respective codes.

Figure 4 Alt text: Four photos of the assembly work-area (a) and its sub-areas (i.e., parts placement area and human's work-area) divided by product: (b) mechanical equipment, (c) tile cutter and (d) diaphragm water pump.

Figure 5 Alt text: Histogram plots showing the distribution of average assembly times for each of the six configurations considered. We observe the following order in terms of increasing average time: 1M, 1HRC, 2M, 2HRC, 3M, 3HRC.

Figure 6 Alt text: Picture showing the position of the six configurations (in ascending order: 1M, 1HRC, 2HRC, 2M, 3HRC, 3M) on the perceived assembly complexity scale obtained through TICS method and their relative distances.

Figure 7 Alt text: Plotting of human-caused process failures divided for the six configurations using boxplot. The median of human-caused process failures is lower in HRC than manual modality for the tile cutter and the pump, while is higher in HRC for the mechanical equipment.

Figure 8 Alt text: Plotting of regression results. We observe an (a) increasing regression line between human-caused process failures and perceived assembly complexity scale and (b) residuals normal Q-Q plot with tails that deviate from normality.

Figure 9 Alt text: Plotting of perceived workload divided for the six configurations using boxplot. The median of perceived workload is lower in HRC than in manual modality for the tile cutter and the pump, while is higher in HRC for the mechanical equipment.

Figure 10 Alt text: Plotting of regression results. We observe an (a) increasing regression line between perceived workload and perceived assembly complexity scale and (b) residuals normal Q-Q plot that shows a normal distribution.

Figure 11 Alt text: Scatter plot of the three manual configurations where x-axis represents TICS-perceived complexity and y-axis the objective complexity from Samy and ElMaraghy H.'s method. The graph shows a linear trend between the two variables.