ASME Journal of Manufacturing Science and Engineering
Online journal at:
https://asmedigitalcollection.asme.org/manufacturingscience

Elisa Verna[1]
Department of Management and Production Engineering,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
Torino 10129, Italy
e-mail: elisa.verna@polito.it

Stefano Puttero
Department of Management and Production Engineering,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
Torino 10129, Italy
e-mail: stefano.puttero@polito.it

Gianfranco Genta
Department of Management and Production Engineering,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
Torino 10129, Italy
e-mail: gianfranco.genta@polito.it

Maurizio Galetto
Department of Management and Production Engineering,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
Torino 10129, Italy
e-mail: maurizio.galetto@polito.it

Exploring the Effects of Perceived Complexity Criteria on Performance Measures of Human–Robot Collaborative Assembly

The use of Human–Robot Collaboration (HRC) in assembly tasks has gained increasing attention in recent years as it allows for the combination of the flexibility and dexterity of human operators with the repeatability of robots, thus meeting the demands of the current market. However, the performance of these collaborative systems is known to be influenced by various factors, including the complexity perceived by operators. This study aimed to investigate the effects of perceived complexity on the performance measures of HRC assembly. An experimental campaign was conducted in which a sample of skilled operators was instructed to perform six different variants of electronic boards and express a complexity assessment based on a set of assembly complexity criteria. Performance measures such as assembly time, in-process defects, quality control times, offline defects, total defects, and human stress response were monitored. The results of the study showed that the perceived complexity had a significant effect on assembly time, in-process and total defects, and human stress response, while no significant effect was found for offline defects and quality control times. Specifically, product variants perceived as more complex resulted in lower performance measures compared to products perceived as less complex. These findings hold important implications for the design and implementation of HRC assembly systems and suggest that perceived complexity should be taken into consideration to increase HRC performance. [DOI: 10.1115/1.4063232]

Keywords: *assembly, inspection and quality control, production systems optimization*

1 Introduction

In today's market, manufacturers are required to produce high-value-added products that meet customer demands at a competitive price, while also complying with sustainability requirements related to environmental and social aspects. As a result, manufacturers must offer a wide range of continuously improved products at competitive prices in order to maintain and increase their market share. Accordingly, balancing high levels of customer adaptation and cost efficiency is crucial in achieving this goal. Research has shown that an increase in product variety not only can lead to a higher market share and sales volume but also increases product complexity and cost [1–3] and requires a flexible manufacturing system that can adapt to changes in product volumes and types [4]. This is especially relevant in the automotive and electronic industries, where frequent changes and an increased number of product variants with more features and functionality are required to meet customer expectations. Managing a large product assortment and assembly conditions can be challenging for manufacturers; however, effectively navigating this complexity can result in a competitive advantage in the industry [5,6].

One approach to achieving mass customization is the use of a traditional manual assembly system, which allows human operators to perform all assembly tasks. However, this approach may result in a decrease in productivity and an increase in costs [7]. On the other hand, automatic assembly systems offer high production rates and cost savings, but they may not be suitable for mass customization [8]. Flexible assembly systems using collaborative robots, or cobots, offer a solution by combining the flexibility of human operators with the precision and accuracy of robots, typically resulting in increased productivity and cost savings [4,9].

The collaboration between humans and cobots, known as human–robot collaboration (HRC), has garnered significant attention in recent years due to the potential benefits and challenges associated with this approach [10]. Previous research in the manual assembly field has shown that assembly complexity and its perception can significantly affect human and process performance [1,11,12]. However, there has been limited research on the impact of perceived assembly complexity on the performance of human–robot collaboration in assembly tasks. Building on these findings,

[1]Corresponding author.
Manuscript received February 22, 2023; final manuscript received August 12, 2023; published online September 1, 2023. Assoc. Editor: Andreas Klink.

the present research aims to extend the investigation by examining the effects of varying product complexity on perceived complexity and assembly performance measures in the context of HRC assembly. This research allows for an understanding of how the perceived complexity of human operators in HRC tasks is influenced by the complexity of the product being assembled.

The main innovative contribution to the field provided by this research is to examine the impact of perceived complexity on several HRC performance measures that encompass the entire manufacturing process. These measures, which include characteristics of the assembly process, the quality control process, and human aspects, are (i) assembly times, (ii) quality control times carried out after the assembly, (iii) in-process defects (catering for errors due to both human and collaborative robots), (iv) offline product defects (i.e., defects detected during offline inspection), (v) total defectiveness (i.e., sum of in-process and offline defects) and (vi) human stress response during assembly. By considering both process performance and human factors, this approach provides valuable insights into the relationship between performance measures and perceived complexity in HRC assembly tasks.

In order to investigate the effects of perceived complexity on HRC performance measures, the study involved the assembly of six variants of electronic boards with different levels of complexity. Skilled operators, assisted by cobots, performed the assembly tasks in a collaborative setup where both humans and cobots worked together in the same workspace [13]. This collaborative configuration is commonly observed in manufacturing environments and facilitates the combination of human dexterity and adaptability with the precision and repeatability of cobots. The adoption of this collaborative mode aimed to investigate the impact of perceived complexity on the performance measures of human–robot collaborative assembly in a real-world context. To ensure a comprehensive analysis of the effects of perceived complexity, a product-centered approach was adopted. The product itself was modified to create different assemblies with varying levels of complexity. This approach is often used in the manufacture of highly customized product variants, where collaboration modes and parameters remain consistent. By focusing on the product and its complexity variations, the study aimed to capture the practical implications of perceived complexity on human–robot collaborative assembly performance measures in an industry-relevant context.

The study's results provide insights into the association between performance measures of human–robot collaboration in assembly tasks and perceived complexity and offer practical implications for designing and implementing high-performing collaborative systems. Furthermore, by considering both process performance and human-related factors, the proposed approach aligns with the goals of sustainable, high-quality, resilient, and human-centric HRC systems within the context of the Industry 5.0 paradigm.

The remainder of the paper is organized as follows. In Sec. 2, the most recent research studies on human–robot collaboration are reviewed. Section 3 presents the experimental details and methods adopted in the present study. In Sec. 4, results are presented and discussed, and the conclusions and future work are outlined in Sec. 5.

2 Literature Review

In recent years, there has been a growing interest in the field of HRC, resulting in a significant increase in research activities and publications. HRC involves the collaboration between humans and robots working together in a shared workspace to perform a task, with each partner contributing their specific skills and abilities [14,15].

The literature on HRC emphasizes the importance of providing technologies that facilitate natural and smooth interactions between humans and robots. Wang et al. [16] highlighted the importance of the communicative interface between robots and humans, to achieve a symbiotic HRC. Inkulu et al. [17] highlighted the prospects and major challenges of HRC, pointing out that human–robot communication modes, such as gestures and speech, enable fluent and immediate interaction, although they still need to be explored in depth.

To date, most research on HRC has focused on safety, communication, and human–robot interaction. Much attention has been given to safety concerns and the development of effective safety measures to support HRC. Indeed, safety is a major concern, especially for robots operating at high speeds and under heavy loads. The introduction of ISO 10218-1:2011 [18] and ISO 10218-2:2011 [19] defined the main hazards that can be encountered when implementing industrial robots in manufacturing environments. In addition, the subsequent ISO/TS 15066:2016 [20] allowed for greater robot's autonomy while working closely with humans. Zanchettin et al. [21] introduced a metric to assess safety in collaborative manufacturing processes. This metric considers human–robot distance, robot type, and operating speed as critical variables affecting safety in HRC. In addition, the sharing of space and time between humans and robots can lead to stress and fatigue issues, which can affect the quality of the output produced and lead to defects in products and processes. Gervasi et al. [22] have developed a conceptual framework for evaluating HRC that includes variables such as mental and physical ergonomics, safety, communication and interaction, team organization, and social acceptance. Advanced adaptive robotic systems are also needed to improve production efficiency.

In manufacturing, concepts such as stress, fatigue, mental workload, and ergonomics have long been addressed [23–25]. Over the years, many tools and methods have been proposed to assess these factors. Self-reporting instruments include the NASA-TLX [26] and the Subjective Workload Assessment Technique (SWAT) [27]. However, these tools have been found to be inappropriate and unreliable in manufacturing environments [28]. Consequently, in recent years, attention has shifted to investigating the impact of objective physiological measures, such as heart rate variability (HRV) and electrodermal activity (EDA), on the operator's state during an HRC task [29–32]. Kulić and Croft [33] investigated how the human physiological state, measured by HRV and EDA, can be affected by the movements of an industrial robot. In this study, proximity and speed were shown to have a significant effect on mental stress. Similarly, Arai et al. [34] evaluated the effect of robot movements, varying operating speed and distance from the operator, on EDA. Kühnlenz et al. [35] studied the effects on humans through HRV and EDA of different trajectory patterns of an industrial robot.

Physical and cognitive aspects are critical factors in the design of HRC tasks [36]. Colim et al. [37] provided guidelines for the design of safe and ergonomic collaborative workstations. In a repetitive and hazardous assembly task, cobots can be used to reduce potential risks to the operator and improve human well-being. However, few studies have investigated the effect of human–robot collaboration on the mental and physical workload perceived by humans. Khalid et al. [38] investigated the safety of HRC systems when using high-load robots, defining potential hazards that include physical and mental strain associated with a collaborative task. Galin and Meshcheryakov [39] analyzed both human and robot-dependent factors that may affect the efficiency of HRC. Among the human factors, emotional and cognitive aspects were found to be critical for HRC efficiency.

Overall, while much attention has been paid to safety, communication, interaction, and human physical and cognitive aspects in HRC, there is a lack of research exploring the impact of task complexity perceived by humans on performance measures. This gap in the literature provides an opportunity for further research to investigate the relationship between assembly complexity and performance measures, both process- and human-related measures, such as production time, defect rates, and human-centered measures, respectively, in HRC settings.

3 Experimental Setup and Methods

3.1 Experimental System Configuration. An experimental campaign involving six expert operators and a single-armed collaborative robot, the UR3e from Universal Robots™, equipped with an OnRobot RG6 gripper with two flexible fingers (see Fig. 1), was designed and carried out. The RG6 gripper, produced by OnRobot™, was selected for its versatility and ability to handle a variety of objects, even of small dimensions. Each operator underwent preliminary training sessions prior to the assembly trials in order to ensure a consistent level of proficiency among the participants and to minimize the potential impact of varying skill levels on the results. These training sessions were designed to familiarize the operators with the assembly process and equipment.

During the experimental trials, each operator assembled six electronic boards (Sec. 3.2) in random order with the support of the UR3e cobot.

Manufacturing process consisted of two phases: (i) assembly phase and (ii) quality control phase. During the assembly phase of each electronic board, the cobot was used to assist operators in assembly operations by passing appropriate components in a predetermined sequence. The parts of the electronic boards were placed in a specific position within the HRC workstation to be picked up by the cobot since the cobot was unable to recognize parts. Future research will focus on the use of visual recognition systems, integrated with machine learning techniques, to enable the cobot to recognize parts. The assembly sequence was determined according to circuit theory [40]. In fact, for the circuit to work, a complete path must exist between the energy source (power) and the lowest energy point (ground). Furthermore, the current always seeks the path of least resistance to earth, and between two possible paths, the current goes through the path of least resistance. This is because the electrical energy within the circuit is dissipated by its components, converting the electrical energy into other forms of energy, such as light, heat, and sound. As a result, the strategy for assembling electronic boards was defined based on the path of the electric current.

During assembly, human operators decided when activating the cobot to pick up the parts and bring them to the storage area by pressing a button near the workstation. The cobot used the MoveL movement for vertical actions, such as picking up and depositing the parts, and the faster MoveJ movement for other actions, such as moving the parts to the storage area. Table 1 shows cobot and gripper parameters used in the HRC assembly.

After the assembly phase, in which electronic board variants were assembled through HRC, a skilled quality controller checked their correct functioning and identified residual defects during the quality control phase. The advantage of using electronic boards is the possibility of verifying their proper functioning by connecting them to the PC and running the code. During the quality inspection, the operator who was in charge of the assembly of the electronic board was asked to fill out a questionnaire on the perceived complexity of the assembly, which will be presented in Sec. 3.4. In detail, at the end of each board variant assembly, the operator evaluated perceived complexity by providing evaluations on some assembly complexity criteria, while at the end of the six assemblies, an overall assessment of the importance of the complexity criteria was given (as per Sec. 3.4). Furthermore, during assembly and quality control phase, data on some performance measures were collected, which will be illustrated in Sec. 3.3.

3.2 Product Assembled. For the assembly of the six electronic boards, the ARDUINO UNO Starter Kit from ARDUINO® was used. This kit includes the physical components necessary for assembling the electronic boards (listed in Table 2) and a software package for programming the microcontrollers. In Table 2, the type and quantity of each component are indicated for each product variant (Variant A–Variant F).

These six products have been selected to cover a wide range of product complexity. According to previous studies [41–43], product variants' total complexity is obtained according to the structural complexity model as a combination of the complexity of product components (C_1), the complexity of assembly connections/liaisons (C_2), and the complexity of product architecture (C_3), according to Eq. (1)

$$C = C_1 + C_2 \cdot C_3 \tag{1}$$

In this study, the Lucas Method [44], widely used in literature and for several industrial applications, was applied to define the complexity of product components and connections (C_1 and C_2). On the other hand, product architecture complexity (C_3) was derived as the average of singular values of the adjacency matrix of the product [41]. In Table 2, the product variants are listed according to increasing complexity C. It is noteworthy that an increase in the number of parts does not necessarily imply an increase in complexity C. As mentioned earlier, the products were assembled in random order by the six operators. Randomizing the order of the six product variants during assembly minimized the impact of learning effects and increased internal validity. This approach controlled for potential confounding variables and prevented observed performance measure differences between product variants from being attributed to increased operator familiarity or experience with the assembly process or equipment. Thus, although the manufacturing sequence was not explicitly controlled, randomization helped minimize its potential impact on the results.

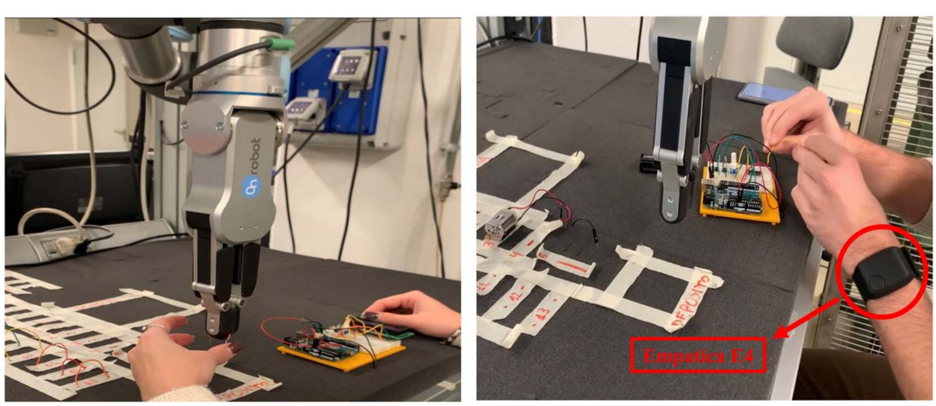

Fig. 1 Collaborative assembly workstation showing the single-armed cobot UR3e (Universal Robots™) with the RG6 gripper (OnRobot™), and product components assembled by an operator wearing the Empatica E4 wristband

Table 1 Cobot and gripper parameters used in the HRC assembly

	Cobot	Gripper
Joint speed (deg/s)	200	–
Joint acceleration (deg/s^2)	200	–
Linear speed (mm/s)	200	–
Linear acceleration (mm/s^2)	200	–
Distance (mm)	–	16
Force (N)	–	80

Figure 2 shows three examples of the six electronic boards assembled with the support of a cobot. The first product, Variant A, is the simplest of the six selected products, Variant C is at medium-level complexity, while the last product, Variant F, is the most complex.

3.3 Data Acquisition. During the manufacturing process, some human and process performance parameters were collected, including physiological data from the operators, the number of total defects (both those occurring during assembly, i.e., in-process defects, and those detected during offline quality control, i.e., offline defects), the assembly time, and the time spent on quality control. The selected performance measures were chosen based on their relevance to the objectives of the study and a thorough literature review that followed the survey proposed by Coronado et al. [36]. While there are many other metrics available for evaluating the performance of collaborative systems, the selected measures were deemed most appropriate for this study due to their widespread use in the manufacturing industry to evaluate the quality of human–robot interaction and collaboration, especially in the context of Industry 5.0, and their ease of monitoring throughout all stages of the production process.

In the first phase of the manufacturing process (assembly phase), information about the assembly time, in-process defects, and stress were collected. On the other hand, in the second phase (quality control phase), information about quality control time and offline defects was collected. Those parameters, plus the total number of defects (sum of in-process and offline defects), are the performance measures depicting the overall manufacturing process.

In the HRC assembly phase, the operator clocked the minutes to complete each electronic board's assembly. The stopwatch started when the cobot picked up the first part and stopped when the operator considered the assembly finished. Even when errors occurred, the stopwatch was never stopped. In the quality control phase, the operator recorded the time in minutes spent on quality control. In this case, the time started when the electronic board reached the quality control station and was stopped when the board worked properly. The stopwatch was never stopped during the quality control phase.

Regarding in-process and offline defects, classification was performed as follows: (i) "Wrong part", i.e., a different component is used instead of the correct one; (ii) "Wrong position", i.e., the component is placed in the wrong position; (iii) "Part not taken", i.e., the cobot fails to pick up the part from the columns; (iv) "Slipped part", i.e., the part slips from the cobot grippers during transport to the operator; (v) "Defective part", i.e., the part is defective and does

Table 2 Characteristics of the six assembled electronic boards

	Variant A	Variant B	Variant C	Variant D	Variant E	Variant F
Long wires	–	1	2	8	9	13
Short wires	1	3	5	3	6	4
Resistors	1	1	4	6	2	2
Pushbuttons	–	2	4	–	2	1
LED	1	1	–	1	–	–
Phototransistor	–	–	–	3	–	–
Potentiometer	–	–	–	–	1	1
Piezo	–	–	1	–	–	–
LCD	–	–	–	–	–	1
Battery snap	–	–	–	–	1	–
DC Motor	–	–	–	–	1	–
H-bridge	–	–	–	–	1	–
No of parts	3	8	16	21	23	22
C_1	1.39	2.87	5.10	6.35	7.25	6.72
C_2	2.98	5.44	13.84	14.58	21.79	26.02
C_3	0.94	0.90	0.90	0.93	0.83	0.84
C	4.20	7.77	17.51	19.95	25.35	28.61

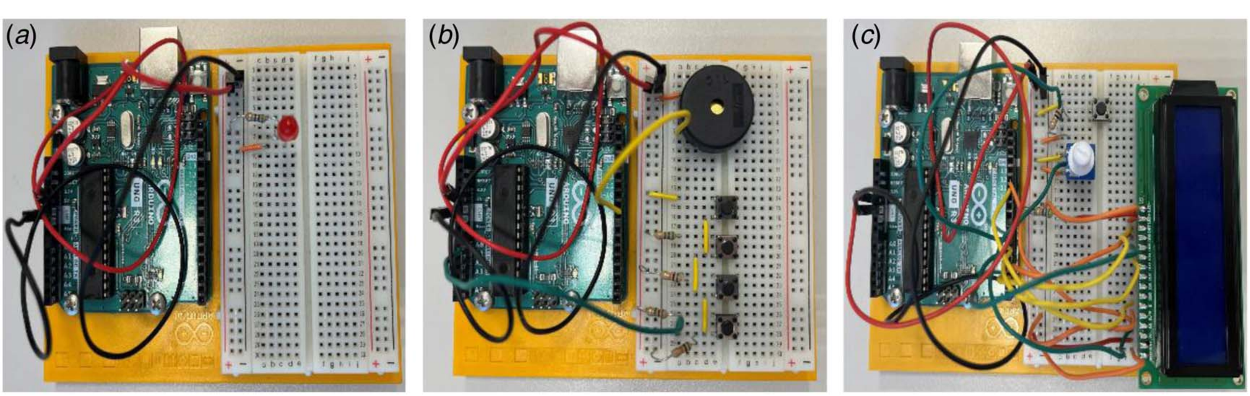

Fig. 2 Example of assembled electronic boards: (*a*) Variant A, (*b*) Variant C, and (*c*) Variant F

not allow the electronic board to function correctly; (vi) "Incorrectly inserted part", i.e., the part is inserted in the correct position but not properly. Obviously, for offline defects, the two categories of defects related to cobot errors ("Part not taken" and "Slipped part") were not present. The assembly operators and the quality control operator collected in-process and offline defect data for each electronic board, indicating the number of defects found for each category.

During the HRC assembly phase, information on the stress level of the operators was collected. Physiological data were measured with the Empatica E4 wristband (Empatica Srl, Milan, Italy), a non-invasive biosensor that records information on ElectroDermal Activity (EDA) at a frequency of 4 Hz (see Fig. 1). EDA is commonly used as an indicator of human stress response, being linked to Skin Conductance Response (SCR) [32]. In detail, continuous signals of tonic and phasic activity constitute the EDA signal. Changes in Skin Conductance Level (SCL) are the best indicator of tonic activity, which is defined as long-term fluctuations in EDA that are not explicitly triggered by external stimuli. Instead, phasic activity describes brief variations in EDA triggered by stimuli typically recognized and presented externally. Skin Conductance Responses (SCRs), i.e., amplitude changes from the SCL, can therefore be detected by examining the phasic activity signal. In this research, the normalized peak amplitude of the SCR was employed as a metric for measuring the stress levels of operators during the HRC assembly of electronic boards. For each operator, the *Human stress response* can be defined as follows:

$$\textit{Human stress response} = \left[\frac{\left(\frac{\sum_{w=1}^{N_P} a_w}{N_P}\right) - a_{\min}}{a_{\max} - a_{\min}}\right] \cdot 100 \tag{2}$$

where a_w is the amplitude of the wth SCR peak, N_P is the total number of SCR peaks during the assembly of a certain product variant, $a_{\min}$ and $a_{\max}$ are, respectively, the minimum and maximum amplitude of SCR peaks obtained during the assembly by each operator.

In this study, the EDA signal was analyzed using the online EDA Explorer software [45]. This software cleans the raw signal of any external noise and identifies peaks in the physiological signal. Figure 3 shows an example of the software output. The trend of the physiological signal (expressed in μS) is the blue line and the green vertical lines represent the peaks identified by the software. In addition, the amplitude of a generic peak (a_w) is shown in red as an example. Furthermore, after assembly, data on perceived complexity were acquired through questionnaires submitted to operators, as described in Sec. 3.4.

3.4 Perceived Complexity Assessment. Complexity, a multi-faceted concept that has been studied extensively and has various definitions and measurements depending on context and research goals, can be assessed objectively, based on inherent task characteristics, or subjectively, considering both task and performer characteristics [46].

This study proposes a complexity assessment framework based on the 16 complexity criteria developed by Falck and Rosenqvist [47] and later adapted for industrial manufacturing sectors [48–50]. The complexity assessments were carried out in collaboration with the company's ergonomist and engineers in the manufacturing engineering department. In order to ensure the easy and quick assembly of the products, Table 3 provides a brief description of each ith criterion ($i = 1, \ldots, 16$), expressed for an easy and fast assembly [50]. For a more detailed description and guidelines for using these criteria in a practical setting, refer to the papers by Falck et al. [50,51].

For each product j, the importance of each criterion i was determined by asking each operator k to assign an importance score (I_{ijk}) using a five-level ordinal scale (see Table 4), based on their perceived relevance for low product complexity. In addition, each operator was asked to indicate the level of agreement (V_{ijk}) with each criterion i in relation to the assembled product j, using the five-level ordinal scale shown in Table 5.

To obtain an estimate of perceived complexity at the individual level, the study combined the operators' ratings of importance and level of agreement with the 16 criteria. However, as the criteria were expressed using linguistic ordinal scales, a systematic method was required to process the data. To this end, the Multi-Expert-Multi-Criteria Decision Making (ME-MCDM) method developed by Yager [52] was adopted as the synthesis approach.

ME-MCDM is a widely used method for aggregating individual operator evaluations to obtain an overall synthetic linguistic value [52]. It combines linguistic information provided for non-equally important criteria using maximum, minimum, and negation operators. The logic behind the ME-MCDM method is that the impact of low-importance criteria on the overall aggregated value should be marginal, while high-important criteria should have a significant impact on the definition of the aggregated evaluation. In the proposed approach, the perceived complexity of the assembly of a product j expressed by the operator k (PC_{jk}) can be calculated using fuzzy logic as follows [53]:

$$PC_{jk} = \mathrm{Min}_i[\mathrm{Max}\{Neg(I_{ijk}),\ V_{ijk}\}] \tag{3}$$

where $Neg(L_x) = L_{t-x+1}$ is the negation of L_x, with L_x the xth level of the scale and t the number of scale levels, i.e., 5 in this case. For instance, $Neg(L_1) = L_5$ and $Neg(L_2) = L_4$.

The rating process for the perceived complexity of a product involves assigning values on a five-point ordinal scale, with the highest level representing low complexity and the lowest level representing high complexity. This scale is based on the criteria listed in Table 3, which are considered to be low-complexity criteria.

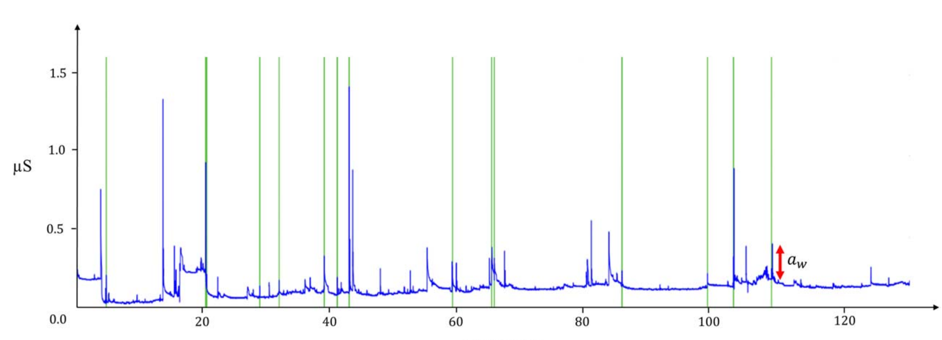

Fig. 3 Example of EDA signal processed with EDA explorer

Table 3 Complexity criteria of assembly, adapted from Falck et al. [47] to suit the electronic platform assembly

Criterion i	Assembly low-complexity criterion	Description
1	Few different ways to perform assembly	Complexity is high if the parts can be assembled/executed correctly in different ways Otherwise, complexity is low if there is a standardized (accepted) way to perform the task
2	Few parts/components and details and few operations	If there are few details to assemble, a small number of operations on the parts, pre-assembly, and module creation (integrated assembly), the complexity is low. Otherwise, complexity is high if there are many details and partial operations
3	Quick and easy operations (no time-demanding operations)	Complexity is low if the solutions are easy and quick to assemble (not time-consuming). Otherwise, if there are time-consuming operations, the complexity is high
4	Clear assembly location of parts/components (immediate understanding of where to place parts within the structure)	If the assembly position of parts and components is clear, the complexity is low; otherwise, it is high
5	Good accessibility to the structure during assembly	If the accessibility to the structure is good (i.e., sufficient for hands/tools), the complexity is low; otherwise, it is high
6	Fully visible operations (operations do not require orientation of the assembly for better visibility)	If the assembly involves visible operations (i.e., in the field of view when looking directly at the structure), the complexity is low; otherwise, it is high
7	Ergonomically easy handling of the structure	If there are good ergonomic conditions, the complexity is low; otherwise, it is high
8	Operator-independent operations that do not require much experience to be performed correctly	If additional training (specialized knowledge) is required beyond the common introductory sessions, then the complexity is high. If the operations do not require additional training, then the complexity is low
9	Operations do not have to be performed in a certain order	If the operations can be performed without following a specific order, that is, they are independent of the order of assembly, the complexity is low. Otherwise, complexity is high if the operations must be performed in a certain order/sequence to complete the assembly correctly
10	Unnecessary intermediate visual checks during assembly to assess the quality and correctness of the structure	If no intermediate checks are required during assembly to assess the quality and correctness of the structure, the complexity is low. Otherwise, complexity is high if visual checks, i.e., careful subjective assessment of quality, are required
11	Operations require little precision, accuracy and attention.	If operations do not require precision and accurate assembly is not necessary, the complexity is low
12	No need for adjustments and corrections (due to errors or inaccuracies) during assembly	The complexity is low if no adjustments are needed due to errors or inaccuracies. Otherwise, the complexity is high
13	Easy to assemble and self-position parts/components that can be controlled in three dimensions: X, Y, and Z	If the surrounding environment varies, where the parts and components will be assembled, or if the detail to be placed depends on the surrounding components, then the complexity is high. Examples of when the geometric environment is varied are: several holes must overlap, components not joined, and components moving relative to each other
14	No detailed instructions are needed and the operator can proceed intuitively	If no detailed instructions are required, i.e., the operator can proceed intuitively to make the assemblies, the complexity is low. Otherwise, the complexity is high
15	The structure does not involve soft and flexible materials (i.e., it is form-resistant)	Complexity is low if the components are rigid and compact and do not change size or deform during assembly. If the structure involves assembling soft and flexible materials, complexity is high
16	There is immediate feedback on correct assembly (e.g., with a clear click and/or compliance with reference points)	Complexity is low if there is immediate feedback of correct assembly, such as through a clear clicking sound and/or adherence to reference points. Otherwise, the complexity is high

Table 4 Scale levels and semantic meanings for assessing product low-complexity criteria importance (I_{ijk})

Scale level	Importance
L_1	Negligible
L_2	Preferable
L_3	Important
L_4	Very important
L_5	Indispensable

Table 6 provides details on the five complexity levels used for individual perceived complexity assessment.

To illustrate how this scoring process works, consider a hypothetical product j, and an operator k, who scores all criteria as L_5—"Indispensable" for importance and L_5—"Totally agree" for agreement. According to the proposed aggregation method, this operator's individual perceived complexity PC_{jk} for product j would be L_5—"Low," meaning that the operator finds the product extremely simple and considers all criteria essential for a simple

Table 5 Scale levels and semantic meanings for assessing agreement degree with low-complexity criteria (V_{ijk})

Scale level	Importance
L_1	Totally disagree
L_2	Disagree
L_3	Relatively agree
L_4	Agree
L_5	Totally agree

Table 6 Scale levels and semantic meanings for the assessment of perceived complexity (PC_{jk})

Scale level	Perceived complexity
L_1	High
L_2	Rather high
L_3	Moderate
L_4	Rather low
L_5	Low

Table 7 Descriptive statistics of performance measures of the six products assembled

Performance measure	Product	Mean	St. dev.	Min	Max
Assembly time (min)	Variant A	1.889	0.627	1.317	2.800
	Variant B	3.928	1.776	1.983	6.967
	Variant C	7.314	1.620	5.833	10.200
	Variant D	9.522	2.238	5.783	12.117
	Variant E	11.719	2.364	8.850	14.800
	Variant F	15.320	4.770	10.430	23.730
Quality control time (min)	Variant A	0.125	0.061	0.000	0.150
	Variant B	0.431	0.436	0.150	1.050
	Variant C	0.769	0.961	0.150	2.083
	Variant D	0.656	0.791	0.150	2.083
	Variant E	1.356	1.875	0.150	4.033
	Variant F	2.308	1.633	0.150	5.183
In-process defects (−)	Variant A	0.000	0.000	0.000	0.000
	Variant B	0.667	0.816	0.000	2.000
	Variant C	1.000	0.894	0.000	2.000
	Variant D	1.833	0.983	0.000	3.000
	Variant E	3.167	1.602	1.000	6.000
	Variant F	3.667	0.816	3.000	5.000
Offline defects (−)	Variant A	0.000	0.000	0.000	0.000
	Variant B	0.333	0.516	0.000	1.000
	Variant C	0.500	0.837	0.000	2.000
	Variant D	0.500	0.548	0.000	1.000
	Variant E	0.500	0.837	0.000	2.000
	Variant F	1.833	1.169	0.000	3.000
Total defects (−)	Variant A	0.000	0.000	0.000	0.000
	Variant B	1.000	0.894	0.000	2.000
	Variant C	1.500	1.378	0.000	3.000
	Variant D	2.333	1.211	0.000	3.000
	Variant E	3.667	1.751	1.000	6.000
	Variant F	5.500	1.049	4.000	7.000
Human stress response (%)	Variant A	0.000	0.000	0.000	0.000
	Variant B	3.180	2.620	0.330	7.350
	Variant C	7.941	2.447	4.021	11.124
	Variant D	12.00	3.390	7.750	16.650
	Variant E	11.99	2.870	9.210	17.310
	Variant F	24.72	5.740	19.840	34.870

Table 8 Classification of in-process (In) and offline (Off) defects for the six assembled products

	Wrong part		Wrong position		Part not taken		Slipped part		Defective part		Incorrectly inserted part	
Product	In	Off	In	Off	In	Off	In	Off	In	Off	In	Off
Variant A	0	0	0	0	0	0	0	0	0	0	0	0
Variant B	0	0	1	1	3	0	0	0	0	0	0	1
Variant C	0	0	5	2	3	0	0	0	0	0	0	1
Variant D	0	0	4	3	4	0	0	0	0	0	3	0
Variant E	0	0	6	3	11	0	2	0	0	0	0	0
Variant F	0	0	11	11	10	0	0	0	0	0	1	0
Total	0	0	27	20	31	0	2	0	0	0	4	2

assembly. Conversely, if the operator rated all criteria importance as L_5—"Indispensable" and the level of agreement as L_1—"Totally disagree", then his individual perceived complexity would be L_1—"High". In this case, the operator considers the product to be extremely complex and considers all criteria to be essential for a simple assembly. In a different scenario, if the operator assigned L_1—"Totally disagree" for agreement degrees, but considers all the criteria to be negligible, resulting in L_1—"Negligible" for importance, the procedure leads to obtain L_5—"Low" for the individual perceived complexity.

Overall, the perceived complexity assessment process involves assigning importance and agreement values to specific criteria, which are then aggregated to determine the individual perceived complexity level of a product assembly.

3.5 Statistical Analysis. The data gathered for the six electronic boards assembled by the 6 operators were collected in a matrix, one line for each product (i.e., 36 rows) with the observed parameters listed in columns. In detail, the parameters related to performance measures recorded in the columns were as follows:

- assembly time,
- quality control time,
- in-process defects,
- offline defects,
- total defects, and
- human stress response (see Eq. (2)).

Furthermore, additional columns were created containing values related to perceived complexity assessment, as follows:

Table 9 Correlation matrix with Spearman correlation coefficients between the agreement degree with the 16 complexity criteria for the six products assembled

Criterion i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2	0.660*															
3	0.446*	0.575*														
4	0.616*	0.576*	0.467*													
5	0.600*	0.489*	0.465*	0.637*												
6	0.559*	0.526*	0.379*	0.730*	0.731*											
7	0.192	0.345*	0.594*	0.420*	0.249	0.400*										
8	0.301	0.205	0.477*	0.446*	0.231	0.295	0.718*									
9	0.240	0.074	−0.080	0.201	0.344*	0.343*	−0.154	−0.138								
10	0.501*	0.265	0.168	0.417*	0.491*	0.577*	0.068	0.173	0.673*							
11	0.085	0.286	0.025	0.259	0.318	0.442*	−0.228	−0.433*	0.500*	0.507*						
12	0.516*	0.307	0.252	0.320	0.487*	0.503*	0.093	0.313	0.614*	0.793*	0.277					
13	0.275	0.276	0.496*	0.535*	0.272	0.416*	0.646*	0.633*	0.011	0.325	−0.019	0.212				
14	−0.113	−0.017	0.187	−0.136	0.139	0.072	0.086	−0.163	−0.294	−0.191	0.225	−0.243	0.015			
15	−0.380*	−0.117	0.004	−0.081	−0.034	−0.022	0.139	−0.161	−0.563*	−0.565*	−0.106	−0.619*	−0.072	0.493*		
16	0.682*	0.486*	0.320	0.620*	0.529*	0.518*	0.437*	0.460*	0.139	0.448*	−0.010	0.441*	0.337*	−0.173	−0.211	

Note: Statistically significant coefficients at 95% confidence level (thus with p-value < 0.05) are asterisked.

- Individual importance evaluations of each of the 16 criteria (as per Table 4).
- Individual agreement degree evaluations of each of the 16 criteria (as per Table 5).
- Individual perceived complexity derived according to Eq. (3).

The primary statistical analysis consisted of calculating the main descriptive statistics for performance measures for each of the six assembled electronic boards (see Table 7 in Sec. 4).

To evaluate if the 16 criteria selected for the analysis compose a suitable set to assess complexity, a pairwise correlation analysis between the evaluations on the agreement degrees provided by operators for each product (V_{ijk}) was performed (see Table 9). Spearman correlation coefficient was adopted being the agreement degrees expressed on the ordinal scale, and the significance of the correlation was assessed by analyzing the p-values [54].

Then, a pairwise correlation analysis was performed to obtain a first indication of the relationships between the agreement degrees of the 16 complexity criteria and performance measures (as shown in Table 10).

Finally, to examine the relations between the individual perceived complexity values derived according to Eq. (3) and the performance measures (see Fig. 4), an Ordinal Logistic Regression (OLR) was adopted, as perceived complexity is an ordinal response defined using a linguistic scale [55]. The OLR is an ordinal regression model that can only be applied to data that meet the proportional odds assumption. The coefficients in the model are estimated using maximum likelihood, computed by using iteratively reweighted least squares [55]. To analyze and interpret the results of the OLR, two steps should be followed [54,56]. First, the p-value and coefficients are examined to analyze the association between the performance measures and individual perceived complexity. The coefficients are useful for determining whether a change in the predictor variable makes any of the events more or less likely, and the odds ratios are provided to compare the odds of two events. Second, the p-values for the Goodness-of-Fit Tests and the measures of association are examined to determine how well the model fits the data. Values of measures of association, including the Somers' D, Goodman and Kruskal indices, and Kendall's index, close to 0 reveal that the model does not have predictive ability. Results of OLR are reported in Tables 11 and 12 and Fig. 5 of Sec. 4. All calculations were performed using the software MINITAB®.

4 Results and Discussion

Descriptive statistics of performance measures considered in this study are listed in Table 7, separately for each electronic board assembled (Variant A–Variant F). An examination of the data reveals that as the complexity of the assembly increases, there is a tendency for performance measures to worsen as a negative impact on assembly time, quality control time, defects rates, and human stress response is encountered. Additionally, as the products move from simple to more complex (i.e., from Variant A—Variant F), the variability associated with performance measures tends to increase, as demonstrated by the increase in standard deviation in Table 7.

Table 8 presents the classification of in-process and offline defects obtained for each of the six assembled product variants, according to the classification provided in Sec. 3.3. An analysis of the data shows that in-process defects are more frequent compared to offline defects. Additionally, within the typology of in-process defects, "Wrong position" and "Part not taken" demonstrate the highest number of defects; whereas for offline defects, "Wrong position" is the most prevalent category. These findings suggest that the manufacturing process is likely facing more issues when the products are in line rather than when they are inspected offline. Furthermore, the frequent occurrence of "Wrong position" for both in-process and offline defects highlights

Table 10 Spearman correlation coefficients between the agreement degree with the 16 complexity criteria for the six products assembled and the performance measures

Criterion i	Assembly time	Quality control time	In-process defects	Offline defects	Total defects	Human stress response
1	−0.354*	−0.073	−0.353*	−0.107	−0.333*	−0.473*
2	−0.663*	−0.183	−0.663*	−0.129	−0.579*	−0.714*
3	−0.533*	−0.184	−0.571*	−0.150	−0.509*	−0.579*
4	−0.252	−0.108	−0.420*	−0.100	−0.366*	−0.552*
5	−0.358*	−0.067	−0.427*	−0.107	−0.389*	−0.503*
6	−0.304	−0.073	−0.302	−0.088	−0.277	−0.465*
7	−0.222	−0.209	−0.489*	−0.157	−0.451*	−0.415*
8	0.027	−0.167	−0.355*	−0.082	−0.332*	−0.225
9	−0.017	−0.019	0.142	−0.044	0.099	−0.013
10	−0.049	0.087	−0.021	0.067	0.013	−0.064
11	−0.310	0.071	0.002	0.051	0.048	−0.186
12	−0.208	−0.159	−0.160	−0.129	−0.190	−0.238
13	−0.033	−0.130	−0.258	−0.122	−0.238	−0.277
14	−0.352*	−0.045	−0.215	−0.120	−0.196	−0.212
15	−0.106	0.007	−0.252	−0.003	−0.193	−0.153
16	−0.248	−0.164	−0.446*	−0.172	−0.440*	−0.435*

Note: Statistically significant coefficients at 95% confidence level (thus with p-value < 0.05) are asterisked.

the need for efficient and accurate placement of parts during the assembly process.

Table 9 displays the results of the pairwise correlation analysis between the evaluations of the agreement degrees with the 16 criteria provided by operators for each product (V_{ijk}). Only the lower triangular part of the matrix is shown in Table 9 because of the symmetry of the matrix. In detail, the Spearman correlation coefficients are reported and those that resulted statistically significant at a 95% confidence level (thus with p-value<0.05) are asterisked. Most statistically significant correlations are positive, showing that as the score on the degree of agreement of one criterion increases, the other also increases. For instance, Criterion 1 is moderately correlated with Criterion 2, as operators agree that a few different ways of performing assembly are associated with few parts/components and details and few operations. On the other hand, only a few of the correlation coefficients in Table 9 are negative. For instance, there is a moderate negative correlation between Criterion 12 and 15 indicating that as operators concur with the fact that the structure is rigid and involves few flexible materials, they perceive a greater need for adjustments. Conversely, fewer adjustments and modifications are required during assembly if the structure incorporates soft and flexible materials. The results presented in Table 9 indicate that the highest correlation coefficient value is 0.731, and there are no correlations that approach a value of 1. Accordingly, it would not be appropriate to eliminate certain criteria as redundant when assessing individual perceived complexity.

Table 10 presents the results of the pairwise correlation analysis conducted to examine the associations between the agreement degree with the 16 complexity criteria and the data pertaining to performance measures. In detail, for each complexity criterion, the evaluations on the agreement degree provided by the six operators for each of the six products (36 values) are correlated with the six performance measures. Spearman correlation coefficients statistically significant at 95% confidence level are asterisked. Almost all the values in Table 10 are negative because as the agreement with the low-complexity criteria increases, operators concur that the product is simple. Therefore, the simpler the product, the less assembly time, quality control time, defects, and stress are. The results indicate a moderate to strong correlation between several of the complexity criteria and performance measures. It should be noted that some criteria do not show a significant correlation with the performance measures (see for example Criteria 9–13 and Criterion 15). However, many of the correlation coefficients have a p-value very close to the significance level.

The correlation coefficients and the asterisks on significant correlation in Table 10 help to identify which criteria have a high degree of correlation with performance measures, providing valuable information to optimize process and design. For example, assembly time, in-process defects, total defects, and human stress response are highly correlated with Criterion 2, indicating that few parts, details, and operations lead to low values of those performance measures. Thus, this information can be used to support decisions towards the design of products or subassemblies with fewer parts, details, and operations in order to decrease assembly time, defects, and human stress.

In addition, Table 10 shows no significant correlations between the agreement degrees with complexity criteria and both quality control time and offline defects. Although these are performance measures of the production process, they appear to be independent of the operators' perception of the process's complexity. This suggests that factors other than the complexity perception of the operators may have more impact on quality control time and offline defects. Further research will be needed to understand the underlying causes of these measures and how they can be improved.

The individual perceived complexity values derived according to Eq. (3) by the ME-MCDM method were obtained by considering both the importance of the 16 criteria and the agreement degrees with the criteria as per Sec. 3.4. The obtained values range from "High" to "Rather low", according to the classification provided in Table 5. Accordingly, no operator considered the assembled products to be extremely simple. Figure 4 illustrates the obtained perceived complexity values and the performance measures for the six product variants. It should be noted that there is a significant amount of variability in the data shown in Fig. 4. This variability is typical of data obtained through self-reported measures such as interviews and questionnaires and should be considered when interpreting the results of this study.

OLR is adopted to model the relationship between quality performances and obtained perceived complexity. In Table 11, the logistic regression table for assembly time is provided [56].

In summary, the results of the analysis presented in Table 11 suggest that there is a statistically significant association between perceived complexity and assembly time since the p-value associated with the predictor is less than the significance level of 5%, and also since the p-value for the test that all slopes are zero is less than 0.05. The odds ratio of 1.19 indicates that operators are more likely to perceive products as more complex as assembly time increases. The positive coefficient associated with assembly

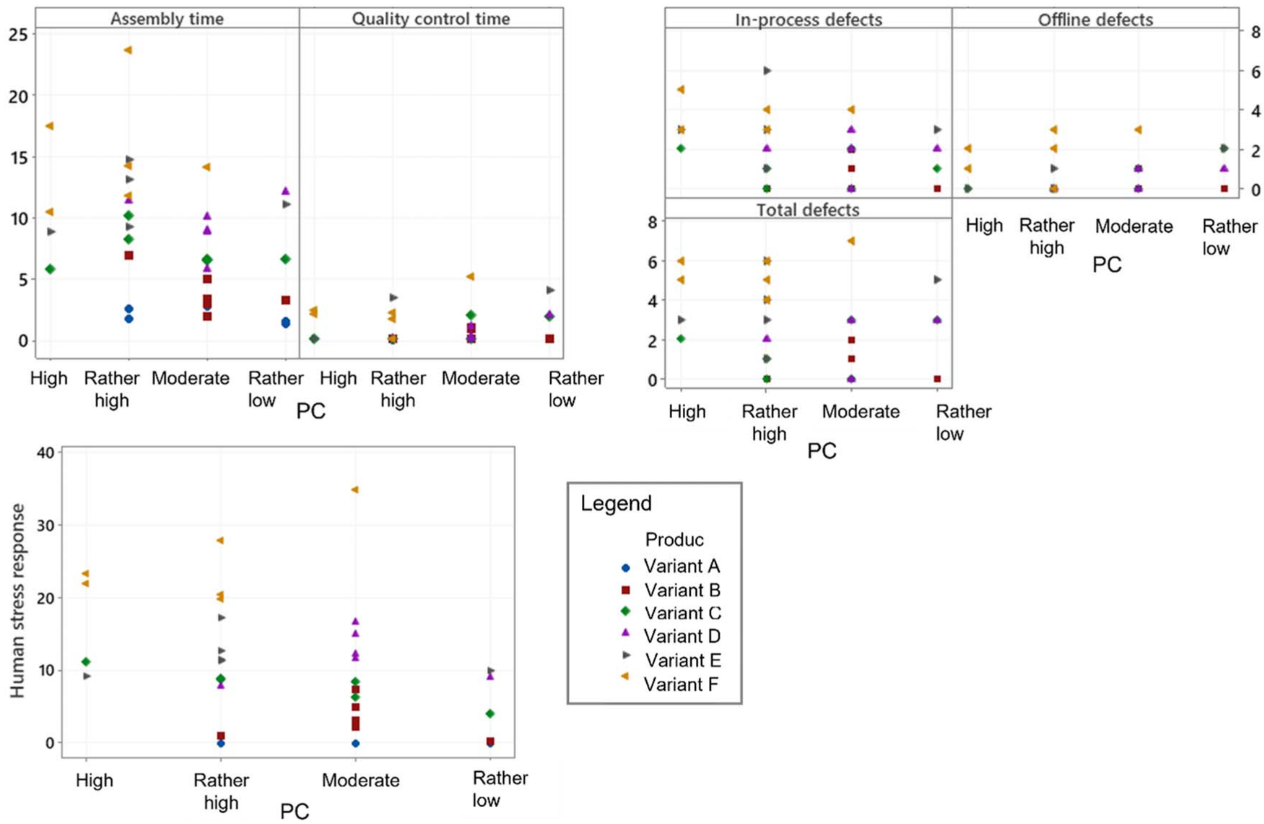

Fig. 4 Scatterplot of individual perceived complexity versus performance measures for the six product variants

Table 11 Logistic regression table for assembly time

Predictor	Coef.	SE Coef.	p-value	Odds ratio	95% confidence interval	
					Lower	Upper
Const(1)	−3.87808	0.924815	0.000			
Const(1)	−1.57885	0.659122	0.017			
Const(3)	0.208728	0.620473	0.737			
Assembly time	0.174226	0.0671240	0.009	1.19	1.04	1.36

Note: Goodness-of-Fit test p-value = 0.905.

time also confirms this result. In addition, the p-value of goodness-of-fit test is greater than 0.05, not providing evidence that the model is inadequate. Overall, this suggests that changes in assembly time are associated with changes in the probabilities of occurrence of the different levels of perceived complexity, as represented in Fig. 5. The data suggests that as assembly time decreases, the probability of the operator perceiving the assembly as “Moderate” or “Rather low” in complexity increases, while an increase in assembly time leads to an increased probability of the assembly being perceived as “High” or “Rather high”. However, the last data point at the maximum assembly time for “Rather high” complexity deviates from this trend; further research is needed to determine the specific cause of this anomaly, as it could be due to operator variability, other factors affecting complexity perception, an outlier data point, or a combination of these factors.

Considering the measures of association reported in Table 12, high values of Somers’ D, Goodman-Kruskal gamma, and Kendall’s tau-a indicate that the model has good predictive ability [56]. These measures are obtained from the number of concordant,

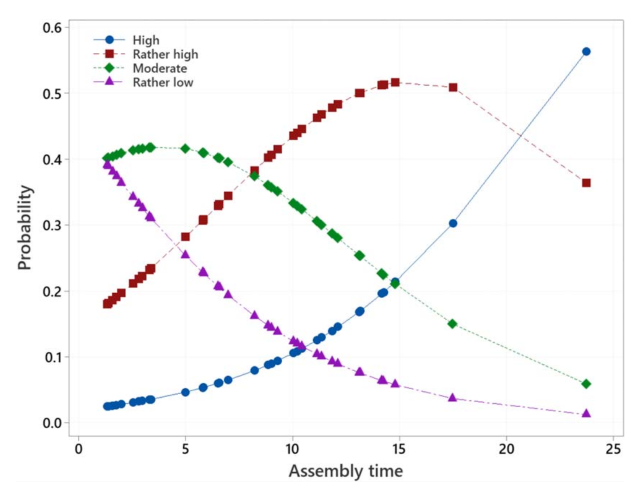

Fig. 5 Probability of occurrence of the levels of individual perceived complexity as a function of assembly time

Table 12 Measures of association between assembly time and predicted probabilities

Pairs	Number	Percent	Summary measures	
Concordant	323	70.4	Somers' D	0.42
Discordant	132	28.8	Goodman-Kruskal Gamma	0.42
Ties	4	0.9	Kendall's Tau-a	0.30
Total	459	100.0		0.42

Table 13 Logistic regression table for In-process defects

Predictor	Coef.	SE Coef.	p-value	Odds ratio	95% confidence interval	
					Lower	Upper
Const(L_1)	−3.19119	0.756213	0.000			
Const(L_2)	−0.958176	0.501900	0.056			
Const(L_3)	0.731504	0.503596	0.146			
In-process defects	0.500009	0.210153	0.017	1.65	1.09	2.49

Note: Goodness-of-Fit test p-value = 0.908.

Table 14 Measures of association between In-process defects and predicted probabilities

Pairs	Number	Percent	Summary measures	
Concordant	263	57.3	Somers' D	0.35
Discordant	103	22.4	Goodman-Kruskal Gamma	0.44
Ties	93	20.3	Kendall's Tau-a	0.25
Total	459	100.0		

Table 15 Logistic regression table for Total defects

Predictor	Coef.	SE Coef.	p-Value	Odds ratio	95% Confidence interval	
					Lower	Upper
Const(1)	−2.78027	0.700359	0.000			
Const(1)	−0.696589	0.490750	0.156			
Const(3)	0.907236	0.511736	0.076			
Total defects	0.258386	0.150963	0.087	1.29	0.96	1.74

Note: Goodness-of-Fit test p-value = 0.493.

Table 16 Measures of association between total defects and predicted probabilities

Pairs	Number	Percent	Summary measures	
Concordant	249	54.2	Somers' D	0.25
Discordant	135	29.4	Goodman-Kruskal Gamma	0.30
Ties	75	16.3	Kendall's Tau-a	0.18
Total	459	100.0		

discordant, and tied pairs, which are calculated by forming all possible pairs of observations (i.e., assembly time values) with the different levels of individual perceived complexity. For the present case study, 459 total pairs were obtained, since four operators perceived the assembly complexity as "High," 13 as "Rather high," 12 as "Moderate," and 7 as "Rather low."

Regarding the other performance measures, the association between perceived complexity and in-process defects, total defects, and human stress response resulted to be statistically

Table 17 Logistic regression table for Human stress response

Predictor	Coef.	SE Coef.	p-Value	Odds ratio	95% Confidence interval	
					Lower	Upper
Const(1)	−3.28926	0.786844	0.000			
Const(1)	−1.11447	0.527668	0.035			
Const(3)	0.602590	0.522103	0.248			
Human stress response	0.0984811	0.0400084	0.014	1.10	1.02	1.19

Note: Goodness-of-Fit test p-value = 0.855.

Table 18 Measures of association between human stress response and predicted probabilities

Pairs	Number	Percent	Summary measures	
Concordant	318	69.3	Somers' D	0.41
Discordant	129	28.1	Goodman-Kruskal Gamma	0.42
Ties	12	2.6	Kendall's Tau-a	0.30
Total	459	100.0		

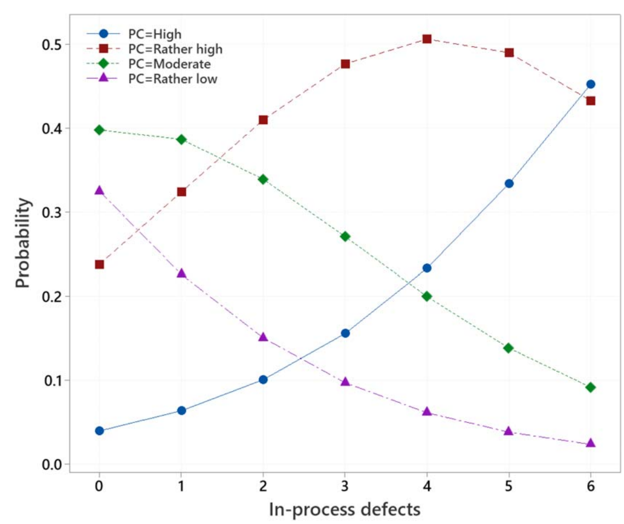

Fig. 6 Probability of occurrence of the levels of individual perceived complexity as a function of In-process defects

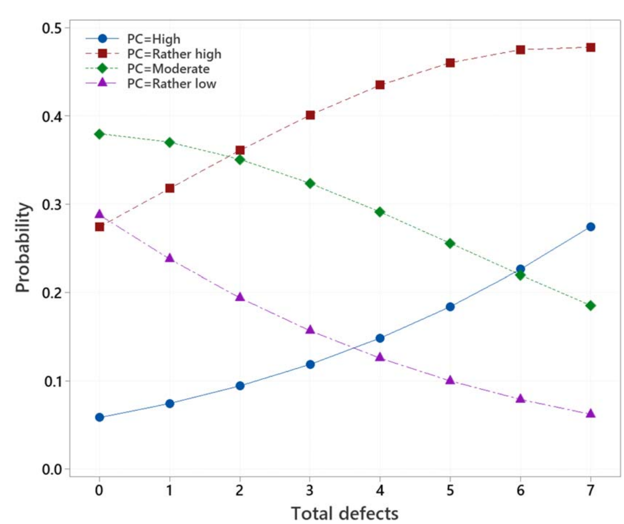

Fig. 7 Probability of occurrence of the levels of individual perceived complexity as a function of total defects

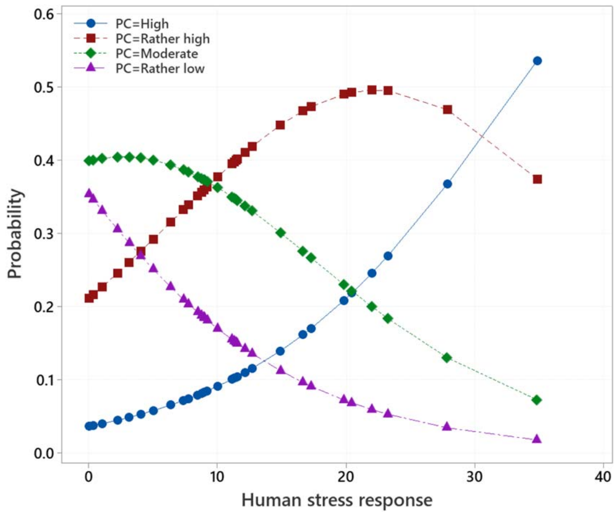

Fig. 8 Probability of occurrence of the levels of individual perceived complexity as a function of human stress response

significant. Tables and figures reporting the results of OLR for such performance measures are given in the Appendix (see Tables 13–18 and Figs. 6–8). Conversely, the association with quality control time and offline defects was found to be not statistically significant, which is consistent with the results of previous correlation analyses (see Table 10).

5 Conclusions

In today's market, manufacturers are required to produce high-value-added products that meet customer demands and expectations at a competitive price while also complying with sustainability requirements. One approach to achieving mass customization is the use of flexible assembly systems that utilize collaborative robots, or "cobots," which can offer increased productivity and cost savings. However, the use of human–robot collaboration in assembly tasks can be impacted by the complexity of the assembly.

This paper focused on the impact of perceived complexity on the performance measures of HRC in assembly tasks. To investigate this issue, the study used a sample of skilled operators to conduct the assembly of six variants of electronic boards with different levels of complexity. Performance measures, including assembly times, quality control times, in-process defects, offline product defects, total defectiveness, and human stress response during assembly, were collected and analyzed. Furthermore, evaluations on the agreement degrees with 16 complexity criteria and their importance provided by the operators for each product were gathered to assess individual perceived complexity. Statistical analysis was conducted on the collected data to quantify the effects of perceived complexity on the HRC performance measures.

The main findings of the present paper are that as complexity perception increases, performance measures tend to worsen, with a negative impact on assembly time, quality control time, in-process defects, and human stress response. Furthermore, for the considered electronic product variants, defects that occurred in-process were more frequent compared to defects detected offline during the quality inspection. The study also showed which complexity criteria are statistically significantly associated with the performance measures, thus providing practical recommendations for engineers to consider when designing processes that focus on reducing perceived complexity and improving overall performance measures. It is important to note that, according to these findings, by reducing perceived complexity, not only the human operators will feel more comfortable with the task but also the process will be more efficient and less error-prone, leading to an increase in productivity and a reduction in costs. Finally, the study highlights that there is no significant association between perceived complexity and the quality control time and the offline defects, indicating that these measures of performance of the production process appear to be independent of the perception that operators have of the complexity of the assembly process. This information is important for engineers to consider in designing and implementing HRC systems as it suggests that a reduction in perceived complexity may not necessarily result in improvements in these specific performance measures. Further studies will need to be conducted to fully understand the underlying reasons and identify potential strategies for improving performance measures related to offline quality control in the HRC assembly process.

The main innovative aspect of this paper is that it considers multiple performance measures linked to both the production and the quality control process, also taking into account human factors such as the operator's perceived stress. By evaluating these measures, this approach allows for a holistic examination of the relationship between perceived complexity and performance, which can provide valuable insights and recommendations for manufacturers to optimize processes and improve performance.

This study has some limitations that should be acknowledged. First, the cobot's involvement in the study was primarily focused on performing pick-and-place operations, which are relatively simple tasks. As a result, the effect of perceived complexity on the cobot's performance and its potential interaction with the perceived complexity of the human operator was not fully explored. Future research should aim to explore different modes of human–robot collaboration, including scenarios where the cobot performs more complex tasks while humans provide support and make key decisions. By considering a broader range of collaboration modes, a more comprehensive understanding of the effects of perceived complexity on HRC performance can be achieved.

Second, the results are based on a specific set of electronic board variants and the subjective concept of perceived complexity may vary among individual operators. Thus, caution is needed when generalizing the findings to other HRC assembly systems. Nonetheless, the study's holistic approach provides practical recommendations for designers and implementers to optimize system performance by considering the subjective perception of complexity by operators. Further research is needed to validate the findings in different contexts and with larger sample sizes to ensure greater statistical power and generalizability.

Additionally, although randomizing the order of the six product variants during assembly helped increase internal validity by minimizing learning effects, the manufacturing sequence was not explicitly controlled. Future research should address this limitation by implementing more systematic control over the manufacturing sequence and by investigating learning effects and their relationship with randomization in more detail.

Finally, based on the derived findings, future work could focus on developing strategies to mitigate the negative effects of perceived complexity on performance measures. One potential approach could be to implement training programs for operators to improve their ability to manage complex product variants. Additionally, improving the design of the assembly process, such as using ergonomic fixtures or improving layout [57], could reduce the complexity of the assembly task and improve performance.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The data sets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

Nomenclature

i = criterions ($i = 1, \ldots, 16$)
j = products ($j = 1, \ldots, 6$)

k = operators ($k = 1, \ldots, 6$)
C = product variants' total complexity
$a_{\max}$ = maximum amplitude of SRC peaks
$a_{\min}$ = minimum amplitude of SRC peaks
a_w = amplitude of the wth SCR peak
C_1 = complexity of product components
C_2 = complexity of assembly connections/liaisons
C_3 = complexity of product architecture
I_{ijk} = importance of criterion i, for product j given by operator k
L_x = xth level of the scale ($x = 1, \ldots, 5$)
N_P = total number of SCR peaks
V_{ijk} = degree of agreement of operator k, for product j on the criterion i
EDA = electrodermal activity
HRC = human–robot collaboration
ME-MCDM = Multi-Expert-Multi-criteria decision making
$Neg(L_x)$ = negation of L_x
OLR = ordinal logistic regression
PC_{jk} = perceived complexity by the operator k for product j
RG6 = gripper produced by OnRobot™
SCL = skin conductance level
SCR = skin conductance response
UR3e = cobot produced by Universal Robots™

References

[1] Falck, A.-C., Örtengren, R., Rosenqvist, M., and Söderberg, R., 2017, "Basic Complexity Criteria and Their Impact on Manual Assembly Quality in Actual Production," Int. J. Ind. Ergon., **58**(Part C), pp. 117–128.
[2] Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022, "Zero Defect Manufacturing: A Self-Adaptive Defect Prediction Model Based on Assembly Complexity," Int. J. Comput. Integr. Manuf., pp. 1–14.
[3] Orfi, N., Terpenny, J., and Sahin-Sariisik, A., 2011, "Harnessing Product Complexity: Step 1—Establishing Product Complexity Dimensions and Indicators," Eng. Econ., **56**(1), pp. 59–79.
[4] Faccio, M., Minto, R., Rosati, G., and Bottin, M., 2020, "The Influence of the Product Characteristics on Human-Robot Collaboration: A Model for the Performance of Collaborative Robotic Assembly," Int. J. Adv. Manuf. Technol., **106**(5–6), pp. 2317–2331.
[5] Rekiek, B., De Lit, P., and Delchambre, A., 2000, "Designing Mixed-Product Assembly Lines," IEEE Trans. Robot. Autom., **16**(3), pp. 268–280.
[6] Zhu, X., Hu, S. J., Koren, Y., and Marin, S. P., 2008, "Modeling of Manufacturing Complexity in Mixed-Model Assembly Lines," ASME. J. Manuf. Sci. Eng.,**130**(5), p. 051013.
[7] Di Pasquale, V., Miranda, S., Neumann, W. P., and Setayesh, A., 2018, "Human Reliability in Manual Assembly Systems: A Systematic Literature Review," Ifac-Papersonline, **51**(11), pp. 675–680.
[8] Hu, S. J., Ko, J., Weyand, L., ElMaraghy, H. A., Lien, T. K., Koren, Y., Bley, H., Chryssolouris, G., Nasr, N., and Shpitalni, M., 2011, "Assembly System Design and Operations for Product Variety," CIRP Ann., **60**(2), pp. 715–733.
[9] Faccio, M., Bottin, M., and Rosati, G., 2019, "Collaborative and Traditional Robotic Assembly: A Comparison Model," Int. J. Adv. Manuf. Technol., **102**(5–8), pp. 1355–1372.
[10] Verna, E., Puttero, S., Genta, G., and Galetto, M., 2022, "Challenges and Opportunities of Collaborative Robots for Quality Control in Manufacturing: Evidences From Research and Industry," Proceedings Book of the 5th International Conference on Quality Engineering and Management, Braga, Portugal, July 14–15, pp. 235–262.
[11] Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022, "Defects-Per-Unit Control Chart for Assembled Products Based on Defect Prediction Models," Int. J. Adv. Manuf. Technol., **119**(5–6), pp. 2835–2846.
[12] Galetto, M., Verna, E., Genta, G., and Franceschini, F., 2020, "Uncertainty Evaluation in the Prediction of Defects and Costs for Quality Inspection Planning in Low-Volume Productions," Int. J. Adv. Manuf. Technol., **108**(11), pp. 3793–3805.
[13] Peshkin, M., and Colgate, J. E., 1999, "Cobots," Ind. Rob., **26**(5), pp. 335–341.
[14] Bauer, A., Wollherr, D., and Buss, M., 2007, "Huma-Robot Collaboration: A Survey," Int. J. Humanoid Rob., **5**(1), pp. 47–66.
[15] Galin, R., and Mamchenko, M., 2021, "Human-Robot Collaboration in the Society of the Future: A Survey on the Challenges and the Barriers," Futuristic Trends in Network and Communication Technologies: Third International Conference, FTNCT 2020, Oct. 14–16, Taganrog, Russia, Revised Selected Papers, Part I 3, Springer, Singapore, pp. 111–122.
[16] Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X. V., Makris, S., and Chryssolouris, G., 2019, "Symbiotic Human-Robot Collaborative Assembly," CIRP Ann., **68**(2), pp. 701–726.
[17] Inkulu, A. K., Bahubalendruni, M. V. A. R., and Dara, A., 2022, "Challenges and Opportunities in Human Robot Collaboration Context of Industry 4.0—A State of the Art Review," Ind. Rob., **49**(2), pp. 226–239.
[18] ISO 10218-1:2011, 2011, *Robots and Robotic Devices. Safety Requirements for Industrial Robots. Robots*, International Organization for Standardization, Genève.
[19] ISO 10218-2:2011, 2011, *Robots and Robotic Devices. Safety Requirements for Industrial Robots. Robot Systems and Integration*, International Organization for Standardization, Genève.
[20] ISO/TS 15066:2016, 2016, *Robots and Robotic Devices. Collaborative Robots*, International Organization for Standardization, Genève.
[21] Zanchettin, A. M., Ceriani, N. M., Rocco, P., Ding, H., and Matthias, B., 2015, "Safety in Human-Robot Collaborative Manufacturing Environments: Metrics and Control," IEEE Trans. Autom. Sci. Eng., **13**(2), pp. 882–893.
[22] Gervasi, R., Mastrogiacomo, L., and Franceschini, F., 2020, "A Conceptual Framework to Evaluate Human-Robot Collaboration," Int. J. Adv. Manuf. Technol., **108**(3), pp. 841–865.
[23] Gawron, V. J., 2008, *Human Performance, Workload, and Situational Awareness Measures Handbook*, CRC Press, Boca Raton, FL.
[24] Young, M. S., Brookhuis, K. A., Wickens, C. D., and Hancock, P. A., 2015, "State of Science: Mental Workload in Ergonomics," Ergonomics, **58**(1), pp. 1–17.
[25] Wickens, C. D., 2008, "Multiple Resources and Mental Workload," Hum. Factors, **50**(3), pp. 449–455.
[26] Hart, S. G., and Staveland, L. E., 1988, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Human Mental Workload*, P. Hancock and N. Meshkati, eds., North Holland, Amsterdam, pp. 139–183.
[27] Reid, G. B., and Nygren, T. E., 1988, "The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload," Adv. Psychol., **52**(C), pp. 185–218.
[28] Marinescu, A. C., Sharples, S., Ritchie, A. C., Sánchez López, T., McDowell, M., and Morvan, H. P., 2018, "Physiological Parameter Response to Variation of Mental Workload," Hum. Factors, **60**(1), pp. 31–56.
[29] Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., and Sharples, S., 2021, "Physiological Indicators of Task Demand, Fatigue, and Cognition in Future Digital Manufacturing Environments," Int. J. Hum. Comput. Stud., **145**, p. 102522.
[30] Bradley, M. M., and Lang, P. J., 1994, "Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential," J. Behav. Ther. Exp. Psychiatry, **25**(1), pp. 49–59.
[31] Gervasi, R., Barravecchia, F., Mastrogiacomo, L., and Franceschini, F., 2023, "Applications of Affective Computing in Human–Robot Interaction: State-of-Art and Challenges for Manufacturing," Proc. Inst. Mech. Eng. Part B J. Eng. Manuf., 237(6–7), pp. 815–832.
[32] Gervasi, R., Aliev, K., Mastrogiacomo, L., and Franceschini, F., 2022, "User Experience and Physiological Response in Human-Robot Collaboration: A Preliminary Investigation," J. Intell. Rob. Syst., **106**(2), p. 36.
[33] Kulic, D., and Croft, E. A., 2007, "Affective State Estimation for Human–Robot Interaction," IEEE Trans. Rob., **23**(5), pp. 991–1000.
[34] Arai, T., Kato, R., and Fujita, M., 2010, "Assessment of Operator Stress Induced by Robot Collaboration in Assembly," CIRP Ann., **59**(1), pp. 5–8.
[35] Kühnlenz, B., Erhart, M., Kainert, M., Wang, Z.-Q., Wilm, J., and Kühnlenz, K., 2018, "Impact of Trajectory Profiles on User Stress in Close Human-Robot Interaction," at-Automatisierungstechnik, **66**(6), pp. 483–491.
[36] Coronado, E., Kiyokawa, T., Ricardez, G. A. G., Ramirez-Alpizar, I. G., Venture, G., and Yamanobe, N., 2022, "Evaluating Quality in Human-Robot Interaction: A Systematic Search and Classification of Performance and Human-Centered Factors, Measures and Metrics Towards an Industry 5.0," J. Manuf. Syst., **63**, pp. 392–410.
[37] Colim, A., Faria, C., Cunha, J., Oliveira, J., Sousa, N., and Rocha, L. A., 2021, "Physical Ergonomic Improvement and Safe Design of an Assembly Workstation Through Collaborative Robotics," Safety, **7**(1), p. 14.
[38] Khalid, A., Kirisci, P., Ghrairi, Z., Thoben, K. D., and Pannek, J., 2017, "Towards Implementing Safety and Security Concepts for Human-Robot Collaboration in the Context of Industry 4.0," Proceedings of the 39th International MATADOR Conference on Advanced Manufacturing, Manchester, UK, July 5–7, pp. 55–63.
[39] Galin, R. R., and Meshcheryakov, R. V., 2020, "Human-Robot Interaction Efficiency and Human-Robot Collaboration," *Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms*, A. G. Kravets, ed., Springer International Publishing, Cham, pp. 55–63.
[40] Zadeh, L., 1962, "From Circuit Theory to System Theory," Proc. IRE, **50**(5), pp. 856–865.
[41] Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022, "Defect Prediction for Assembled Products: A Novel Model Based on the Structural Complexity Paradigm," Int. J. Adv. Manuf. Technol., **120**(5–6), pp. 3405–3426.
[42] Sinha, K., 2014, "Structural Complexity and Its Implications for Design of Cyber-Physical Systems," Ph.D. dissertation, Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA.
[43] Alkan, B., and Harrison, R., 2019, "A Virtual Engineering Based Approach to Verify Structural Complexity of Component-Based Automation Systems in Early Design Phase," J. Manuf. Syst., **53**, pp. 18–31.
[44] Lucas, S. M., and Reynolds, T. J., 2003, "Learning DFA: Evolution Versus Evidence Driven State Merging," The 2003 Congress on Evolutionary Computation, 2003. CEC '03, IEEE, pp. 351–358.

[45] Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., and Picard, R., 2015, "Automatic Identification of Artifacts in Electrodermal Activity Data," Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, Aug. 25–29, pp. 1934–1937.
[46] Liu, P., and Li, Z., 2012, "Task Complexity: A Review and Conceptualization Framework," Int. J. Ind. Ergon., **42**(6), pp. 553–568.
[47] Falck, A. C., and Rosenqvist, M., 2012, "What Are the Obstacles and Needs of Proactive Ergonomics Measures at Early Product Development Stages?—An Interview Study in Five Swedish Companies," Int. J. Ind. Ergon., **42**(5), pp. 406–415.
[48] Falck, A.-C., Örtengren, R., and Rosenqvist, M., 2014, "Assembly Failures and Action Cost in Relation to Complexity Level and Assembly Ergonomics in Manual Assembly (Part 2)," Int. J. Ind. Ergon., **44**(3), pp. 455–459.
[49] Alkan, B., 2019, "An Experimental Investigation on the Relationship Between Perceived Assembly Complexity and Product Design Complexity," Int. J. Interact. Des. Manuf., **13**(3), pp. 1145–1157.
[50] Falck, A.-C., Örtengren, R., Rosenqvist, M., and Söderberg, R., 2017, "Proactive Assessment of Basic Complexity in Manual Assembly: Development of a Tool to Predict and Control Operator-Induced Quality Errors," Int. J. Prod. Res., **55**(15), pp. 4248–4260.
[51] Falck, A.-C., Örtengren, R., Rosenqvist, M., and Söderberg, R., 2016, "Criteria for Assessment of Basic Manual Assembly Complexity," Procedia CIRP, **44**, pp. 424–428.
[52] Yager, R. R., 1993, "Non-Numeric Multi-Criteria Multi-Person Decision Making," Group Decis. Negot., **2**(1), pp. 81–93.
[53] 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," Res. Eng. Des., **34**(3), pp. 301–325.
[54] Agresti, A., 2003, *Categorical Data Analysis*, John Wiley & Sons, Hoboken, NJ.
[55] McCullagh, P., 1980, "Regression Models for Ordinal Data," J. R. Stat. Soc. Ser. B, **42**(2), pp. 109–127.
[56] Powers, D., and Xie, Y., 2008, *Statistical Methods for Categorical Data Analysis*, Emerald Group Publishing, Bingley, UK.
[57] ElMaraghy, W., ElMaraghy, H., Tomiyama, T., and Monostori, L., 2012, "Complexity in Engineering Design and Manufacturing," CIRP Ann., **61**(2), pp. 793–814.