

A Novel Diagnostic Tool for Human-Centric Quality Monitoring in Human-Robot Collaboration Manufacturing

*Original*

A Novel Diagnostic Tool for Human-Centric Quality Monitoring in Human-Robot Collaboration Manufacturing / Verna, E., Puttero, S., Genta, G., Galetto, M.. - In: JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING. - ISSN 1087-1357. - ELETTRONICO. - 145:12(2023). [10.1115/1.4063284]

*Availability:*

This version is available at: 11583/2981357 since: 2023-09-11T13:37:31Z

*Publisher:*

ASME Digital Collection

*Published*

DOI:10.1115/1.4063284

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

ASME postprint/Author's accepted manuscript

© ASME. This is the author' version of the following article: A Novel Diagnostic Tool for Human-Centric Quality Monitoring in Human-Robot Collaboration Manufacturing / Verna, Elisa; Puttero, Stefano; Genta, Gianfranco; Galetto, Maurizio published in : JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING, 2023, <http://dx.doi.org/10.1115/1.4063284>. This author's accepted manuscript is made available under

(Article begins on next page)

# 1 A Novel Diagnostic Tool for Human-Centric 2 Quality Monitoring in Human-Robot 3 Collaboration Manufacturing

4 **Elisa Verna<sup>1</sup>**

5 Department of Management and Production Engineering, Politecnico di Torino  
6 Corso Duca degli Abruzzi 24, 10129 Torino, Italy  
7 elisa.verna@polito.it

8  
9 **Stefano Puttero**

10 Department of Management and Production Engineering, Politecnico di Torino  
11 Corso Duca degli Abruzzi 24, 10129 Torino, Italy  
12 stefano.puttero@polito.it

13  
14 **Gianfranco Genta**

15 Department of Management and Production Engineering, Politecnico di Torino  
16 Corso Duca degli Abruzzi 24, 10129 Torino, Italy  
17 gianfranco.genta@polito.it

18  
19 **Maurizio Galetto**

20 Department of Management and Production Engineering, Politecnico di Torino  
21 Corso Duca degli Abruzzi 24, 10129 Torino, Italy  
22 maurizio.galetto@polito.it

## 23 24 25 **ABSTRACT**

26 *The manufacturing industry is currently facing an increasing demand for customized products, leading to a*  
27 *shift from mass production to mass customization. As a result, operators are required to produce multiple*  
28 *product variants with varying complexity levels while maintaining high-quality standards. Further, in line*  
29 *with the human-centered paradigm of Industry 5.0, ensuring the well-being of workers is equally important*  
30 *as production quality. This paper proposes a novel tool, the "Human-Robot Collaboration Quality and Well-*  
31 *Being Assessment Tool" (HRC-QWAT), which combines the analysis of overall defects generated during*  
32 *product variant manufacturing with the evaluation of human well-being in terms of stress response. The*  
33 *HRC-QWAT enables the evaluation and monitoring of human-robot collaboration systems during product*

---

<sup>1</sup> Corresponding author: Tel: +39 0110907236, e-mail: elisa.verna@polito.it.

34 *variant production from a broader standpoint. A case study of collaborative human-robot assembly is used*  
35 *to demonstrate the applicability of the proposed approach. The results suggest that the HRC-QWAT can*  
36 *evaluate both production quality and human well-being, providing a useful tool for companies to monitor*  
37 *and improve their manufacturing processes. Overall, this paper contributes to developing a human-centric*  
38 *approach to quality monitoring in the context of human-robot collaborative manufacturing.*

39

## 40 **1. INTRODUCTION**

41 Mass production has long been the standard in manufacturing, allowing efficient  
42 use of available resources and a corresponding reduction in production costs. In recent  
43 years, however, there has been a shift towards mass customization, an approach to  
44 production that allows products to be individually tailored to meet each customer's  
45 specific needs and preferences [1]. Several factors have driven this shift, including  
46 technological advances, increased demand for customized products, and growing  
47 awareness of mass production's environmental and social impacts [2]. As a result,  
48 manufacturers are increasingly turning to mass customization as a way to remain  
49 competitive in the marketplace and meet the evolving needs of their customers, thus  
50 representing a significant paradigm shift in how goods are produced and consumed.

51 While greater product variety can increase market share and sales volumes, it also  
52 increases product complexity and costs [1]. As a result, mass customization requires a  
53 flexible production system that can adapt to product volume and type variations. An  
54 effective approach to mass customization is the use of collaborative robots (also called  
55 cobots) in what is known as Human-Robot Collaboration (HRC) [3]. This approach  
56 combines the flexibility and versatility of human operators with the precision of cobots,

57 creating a flexible system capable of assembling different product variants in the same  
58 workstation [4].

59 Interest in HRC has grown with the development of Industry 4.0 and is becoming  
60 increasingly important with the emergence of Industry 5.0. Indeed, the main goal of  
61 Industry 5.0 is to put human well-being at the center of production systems in order to  
62 provide sustainable prosperity for long-term development [5–7]. The technologies and  
63 processes used in Industry 5.0 are designed to enhance the capabilities of human  
64 operators, as in the case of cobots. This approach departs from traditional production  
65 methods, which prioritize efficiency and automation over the satisfaction of human  
66 workers. This shift towards human-centred production should lead to improved  
67 productivity, greater job satisfaction and a more sustainable manufacturing sector  
68 towards mass customization.

69 While the benefits of HRC are clear, there is a lack of comprehensive tools that  
70 can assess and monitor the quality and well-being aspects of these systems. Therefore, a  
71 need exists for an evaluation tool that not only considers production quality but also  
72 considers the well-being of the human operator. With the growing emphasis on human-  
73 centered production in Industry 5.0, such an assessment tool becomes increasingly  
74 crucial.

75 To address the lack of comprehensive tools for assessing and monitoring HRC  
76 systems taking into account both process performance and human-centered  
77 performance for highly customized product variants, this research aims to answer the  
78 following research question: How can a diagnostic tool be developed to assess the

79 performance of HRC systems considering the specificities of highly customized product  
80 variants?

81 In response to this research question, this paper introduces a novel tool, the  
82 "Human-Robot Collaboration Quality and Well-Being Assessment Tool" (HRC-QWAT). This  
83 tool integrates two indicators, a product and process quality indicator and a human stress  
84 response indicator, to assess and monitor the quality of an HRC system, specifically  
85 tailored to accommodate the unique challenges and variations associated with highly  
86 customized product variants. Unlike previous studies that focused on individual  
87 measurements and indicators, the HRC-QWAT incorporates multiple dimensions of  
88 evaluation, encompassing product and process quality as well as human well-being. This  
89 comprehensive approach enables a holistic assessment of the collaborative assembly  
90 process, capturing the interplay between quality outcomes and the well-being of human  
91 operators.

92 The versatility and adaptability of the HRC-QWAT are demonstrated by its  
93 applicability to single-variant and small-batch variant production scenarios, catering to  
94 diverse production contexts. Thus, the HRC-QWAT fills a critical gap in assessing and  
95 monitoring HRC systems, providing practitioners and researchers with a valuable tool for  
96 evaluating, diagnosing, and optimizing HRC systems in the context of diverse and  
97 customized product variants.

98 To show the practical implementation of the HRC-QWAT, a real-life case study was  
99 conducted involving the assembly of electronic board variants using a human-robot  
100 collaborative assembly system. The methodology consists of two main phases: (1) the

101 realization phase, in which the HRC-QWAT is constructed by collecting historical  
102 experimental data and developing a model that relates the two performance measures,  
103 i.e. total defects (cobot-related and human-related errors) and human stress response, to  
104 represent the overall quality of the system; and (2) the use phase, in which the HRC-QWAT  
105 is used as a reference for predicting future products/batches and identifying critical  
106 products in terms of defects and human stress response. The HRC-QWAT can be used to  
107 identify critical production scenarios and implement necessary corrective actions to  
108 maintain the desired quality level while taking into account the well-being of human  
109 operators, thus advancing human-centered production practices within the framework of  
110 Industry 5.0.

111 The remaining paper is structured as follows. Section 2 summarizes the main  
112 studies in the field of quality in HRC. Section 3 presents the HRC assembly system used as  
113 a case study. Section 4 illustrates the complexity assessment of the product HRC  
114 assemblies. Section 5 presents the data collected on product and process quality and their  
115 relationship with assembly complexity. Section 6 illustrates data on human well-being and  
116 discusses the relationship with assembly complexity of product variants. Section 7  
117 presents the novel diagnostic tool called HRC-QWAT, which shows the potential for single  
118 variant and small-batch production. Finally, Section 8 concludes the paper.

119  
120 **2. LITERATURE REVIEW**

121 HRC is a rapidly developing field with promising applications in service, social, and  
122 industrial contexts. When designing and implementing an HRC system, evaluating its  
123 quality is crucial to ensure that the system meets individual, collective, and production

124 needs or objectives. From an engineering point of view, quality refers to the degree to  
125 which a system, product, service, or process conforms to specified requirements and  
126 conditions [6]. Quality models, such as conceptual or definition models, are commonly  
127 used in engineering to support, evaluate, and manage quality [8]. These models outline a  
128 set of quality attributes and their potential interrelationships and serve as a guide for  
129 selecting relevant factors for experimental validation of applications, services or systems.

130 However, efforts to identify and classify factors, measures and metrics that  
131 describe quality in the HRC field are still rare, especially from a human-centred  
132 perspective [6,9]. The industrial perspective can be categorized into two interests:  
133 performance-centred and human-centred. The former considers robots as a means to  
134 optimize the production process, often involving full automation and substituting human  
135 workers with machines, while the latter aims to improve human well-being by respecting  
136 their role, needs, job, talents, and rights [10–12]. Consequently, there is a trade-off  
137 between optimising the production process and optimising the well-being of the  
138 operators, which requires the use of performance measures specific to the collaborative  
139 environment.

140 According to [6], performance measures for HRC are variables that can be  
141 obtained from physical measurements or an aggregate of facts to assess the current or  
142 final state of the human, robot, process, or interaction. These measures can be grouped  
143 into four categories:

- 144 • Time behavior measures indicate the response and processing times required to  
145 perform functions or complete tasks.

- 146       • Process measures are an aggregation of facts related to task completion,  
147       workspace design, safety, or product quality.
- 148       • Physiological measures are obtained from body measures, such as heart rate, to  
149       understand the current state of the human.
- 150       • Human-Robot physical measures are obtained from sensors that indicate the  
151       current state of the interaction, such as the distance between the human and the  
152       robot.

153       Moreover, performance metrics for HRC can be defined as a combination of direct  
154       measures used to express a rate, average or input/output relationship [6]. Efficiency and  
155       effectiveness are considered the main attributes used to evaluate such performance.  
156       Efficiency metrics assess the use of resources, i.e. the input/output ratio. On the other  
157       hand, effectiveness metrics assess the accuracy and completeness of the achievement of  
158       specific objectives, measuring the relationship between actual and expected results.  
159       These metrics assess whether HRC systems are "doing things right" and "doing the right  
160       things", respectively.

161       As far as the human-centred perspective is concerned, quality factors that have  
162       received more attention in the robotics literature are safety [13], trust [14], attitudes and  
163       acceptance [15], mental and physical workload [16,17], situation awareness and mental  
164       models [18,19], emotional responses [20,21] and anxiety [22].

165       Additionally, the review paper [6] identifies seven emergent research topics that  
166       could have a significant impact on future Industry 5.0 applications, including (i) non-  
167       invasive monitoring and online analysis of human factors, (ii) individualized HRC, (iii)

168 transparent robotic systems, (iv) fluency, (v) adaptive workload systems, (vi) privacy in  
169 data-driven HRC, and (vii) benchmarks.

170       Regarding point (i), the Industry 5.0 paradigm aims to optimize human well-being  
171 through human-centred smart environments. However, most tools for assessing human  
172 factors in HRC require offline or intrusive techniques. Creating accurate, non-invasive, and  
173 online ergonomic assessment tools that require short preparation represents a relevant  
174 challenge in HRC for manufacturing settings [23,24]. One of the most widely used tools  
175 for this purpose is the digital twin, which allows the human comfort and flexibility of the  
176 cobot to be improved in a non-intrusive way [25]. Several DT applications have already  
177 been implemented in the area of collaborative assembly and disassembly [26].

178       Concerning point (ii), applications enabling collaborations between humans and  
179 robots are generally short and static for practical reasons [27]. However, individualized  
180 machine collaboration is essential for Industry 5.0. Nowadays, various technologies have  
181 been identified that enable machine collaboration, such as human action recognition,  
182 intention prediction, augmented, virtual or mixed reality, exoskeletons, and collaborative  
183 robots. Personalized HRI systems can continuously collect and process personal and  
184 physiological data, adapt to individuals' needs and preferences, and maintain long-term  
185 interactions [27,28]. Hedonics factors, which mostly focus on individual goals, require  
186 more research attention on applications for Industry 5.0 [29]. Additionally, human-  
187 centred initiatives need to consider technologies enabling job satisfaction, work-life  
188 balance, and up-skilling and re-skilling of workers [12].

189           Regarding point (iii), Industry 4.0 applications use black-box AI to enhance  
190 autonomy, while Industry 5.0 requires transparent AI that interacts with humans. In HRC,  
191 this transparency includes predictability, legibility, and explainability. Legibility enables  
192 observers to quickly infer correct goals, while predictability matches expectations.  
193 Creating legible trajectories is a broad open issue. Multimodal systems for anticipating  
194 human actions face high-dimensional data, which dimensionality reduction techniques  
195 can address. AI aims to explain robot behavior to users and improve trust and situational  
196 awareness, but challenges include creating methods for generating explanations and  
197 evaluating their effectiveness [30,31].

198           With regard to point (iv), fluency is not considered a metric but rather a quality of  
199 interaction in HRC, as described in [32]. In HRC environment, fluency refers to the  
200 seamless interaction between humans and robots. It involves a high level of coordination,  
201 well-synchronized joint activities with precise and efficient timing, and dynamic  
202 adaptation of plans and actions. However, fluency is still a relatively new concept in HRC  
203 research, and proposed metrics for fluency are often task-specific [32]. Recent studies,  
204 such as those of Hoffman [32], have categorized fluency metrics as subjective or objective.  
205 However, due to the somewhat vague and ephemeral nature of fluency, it remains a topic  
206 of discussion in the robotics community, with further research needed to fully understand  
207 the factors affecting fluency and to design metrics that can assess it in various  
208 collaborative settings.

209           Concerning point (v), real-time workload assessment algorithms using  
210 physiological measures, such as heart rate, electrodermal activity and skin temperature,

211 can accurately estimate optimal workload levels in humans [17,33]. This information can  
212 be used to improve task performance, reduce errors, and prevent accidents by changing  
213 interaction mediums, level of autonomy, and reallocating tasks and responsibilities  
214 between humans and robots [33]. Such systems are called adaptive workload or adaptive  
215 teaming systems [17]. The use of these algorithms in various human-robot teaming  
216 scenarios remains an open challenge [34].

217       Regarding point (vi), data-driven technologies like big data, machine learning,  
218 cloud computing, and IoT can enhance production performance and human working  
219 conditions. However, Industry 4.0 has largely overlooked the human factor and the  
220 privacy issues arising from the collection, storage and processing of personal data that  
221 these technologies entail [35,36]. In human-centred manufacturing, privacy efforts must  
222 focus on protecting workers' personal information and ensuring data security [35].  
223 Cybersecurity assessment criteria for HRI in automobile manufacturing have been  
224 proposed [37], but comprehensive metrics are needed for HRI and HRC [37].

225       Finally, with respect to point (vii), international robotics competitions have  
226 become a valuable tool for evaluating the performance of robotics systems, providing a  
227 form of reproducibility and enabling the evaluation of non-competing systems. Although  
228 the scoring mechanism tends to hide the underlying characteristics of the system,  
229 competitions allow systems to be compared by linking relevant metrics to the score and  
230 explaining which aspects influenced the score and in what way. Typically, the score is  
231 based on objective task completion (e.g., image classification accuracy), with few  
232 competitions evaluating safety in HRI. However, there is a shift towards more human-

233 centred objective evaluations, exemplified by the safety score in the Future Convenience  
234 Store Challenge [38,39].

235 Based on the literature review, it is evident that many approaches proposed in the  
236 field of HRC have a performance-centred perspective, which fails to consider the full  
237 potential of HRC applications. Towards a human-centred society and industry, HRC  
238 researchers should broaden their perspective beyond mere task fulfilment and adopt  
239 holistic approaches that enable robotic systems to achieve both collective and individual  
240 goals. In line with this viewpoint and to help address the challenges identified in the seven  
241 emerging research topics, the "Human-Robot Collaboration Quality and Well-Being  
242 Assessment Tool" (HRC-QWAT) has been proposed. In detail, the proposed tool can  
243 respond to the challenges mentioned above as follows:

244 (i) Non-invasive monitoring and online analysis of human factors: HRC-QWAT allows for  
245 real-time, non-invasive monitoring of human operators' stress levels and well-being  
246 through the integration of wearable devices and sensors. This real-time evaluation  
247 ensures a prompt intervention to reduce stress levels, fostering a more efficient and  
248 balanced working environment.

249 (ii) Individualized HRC: The tool offers the possibility of individualized HRC by considering  
250 the unique physiological responses of each worker. This personalized approach promotes  
251 a more efficient and harmonious human-robot interaction, potentially leading to  
252 improved productivity and well-being.

253 (iii) Transparent robotic systems: Transparency is facilitated as the tool evaluates the  
254 collaborative process based on clear performance indicators and stress responses. These

255 evaluations can be shared with human operators, fostering an understanding of the robot  
256 function and promoting trust and collaboration.

257 (iv) Fluency: By measuring the quality of the collaboration through multiple performance  
258 indicators, the HRC-QWAT contributes to assessing the fluency of the human-robot  
259 interaction and collaboration. This analysis promotes the optimization of joint actions and  
260 the creation of more fluid and synchronized interactions.

261 (v) Adaptive workload systems: The HRC-QWAT real-time monitoring of human stress  
262 responses can inform adaptive systems. By detecting stress or overwork, the system can  
263 automatically adjust the workload distribution between human and robot, improving  
264 efficiency and reducing the risk of human error or health implications.

265 (vi) Privacy in data-driven HRC: While HRC-QWAT uses data-driven methods for  
266 evaluation, it is designed with the utmost respect for privacy standards. Personal and  
267 sensitive data are strictly used for the intended purpose of enhancing human-robot  
268 interaction and are safeguarded according to the highest security protocols.

269 (vii) Benchmarks: The proposed tool also serves as a benchmarking instrument for HRC in  
270 different scenarios. By providing comprehensive metrics on both the performance of the  
271 collaborative process and the well-being of the human operator, the HRC-QWAT offers a  
272 valuable standard against which different collaboration setups can be compared.

273 Accordingly, the HRC-QWAT serves as a comprehensive tool, specifically  
274 addressing the identified challenges in HRC, thereby offering a strategic instrument for  
275 human-centered Industry 5.0.

276

277 **3. HRC ASSEMBLY SYSTEM**

278 An experimental campaign is conducted to assemble six different variants of  
279 electronic boards (from variant V1 to variant V6) using the ARDUINO UNO starter kit  
280 (ARDUINO®). The choice of electronic boards is based on the fact that, by using the same  
281 components, highly customized products with varying levels of complexity can be  
282 assembled (as will be discussed in the next Section 4). Moreover, these boards allow real-  
283 time verification of the correct functioning of the products, i.e., their proper assembly.

284 The ARDUINO UNO starter kit is composed of: (i) the components that are  
285 assembled to make up the various boards listed in Table 1 (e.g., the jumper wires that  
286 carry current between the various components); (ii) the microcontroller, i.e., a small  
287 computer that enables the circuits to function; and (iii) the breadboard, i.e., a board on  
288 which the actual circuit can be built. The breadboard consists of rows and columns of  
289 holes that conduct electricity through thin metal connectors under the plastic screen,  
290 allowing the circuit components to be connected. The ARDUINO UNO Breadboard is  
291 defined as 'seamless' as the components do not need to be welded but simply inserted  
292 into the holes. Fig. 1(a) shows an example of an assembled electronic board (variant V3),  
293 while Fig. 1(b) displays the product circuit diagram.

294 Each of the selected products has a varying number of components, which are  
295 connected to the breadboard. As outlined in Section 4, the six electronic boards were  
296 chosen to span a broad range of assembly complexity. Table 1 indicates the type and  
297 number of components required for each of the six electronic board variants (V1 – V6).

298           The assembly of the six electronic board variants was conducted using a Universal  
299 Robots™ UR3e cobot, as depicted in Fig. 2. The boards were assembled using an  
300 OnRobot™ RG6 gripper, a versatile gripper capable of handling small objects and a range  
301 of other objects. Six skilled operators, proficient in electronics and electrical engineering,  
302 were involved in the assembly process of all six electronic boards, following a random  
303 order to prevent any learning effects. During the preliminary stages, each operator  
304 underwent training sessions to ensure consistent proficiency among the participants and  
305 minimize the potential impact of varying skill levels on the results. These training sessions  
306 were carefully designed to familiarize the operators with the assembly process and  
307 equipment, allowing them to develop a solid understanding of the tasks involved in the  
308 HRC assembly. Table 2 provides an overview of the participants' characteristics, including  
309 relevant information such as age, gender, domain knowledge of HRC, and domain  
310 knowledge of assembly tasks. The inclusion of skilled operators with expertise in  
311 electronics and electrical engineering helped ensure that the participants were familiar  
312 with the intricacies of electronic board assembly and could contribute effectively to the  
313 HRC trials.

314           In the assembly phase, the cobot handed over the required components to the  
315 operator, who assembled the electronic boards in a predetermined order, defined based  
316 on circuit theory [40]. The operator completely controlled the logistic tasks by activating  
317 the cobot using a button. After the assembly was completed, an experienced external  
318 operator (who was not involved in the assembly) conducted an offline quality control  
319 check to identify any defects in the final product. Data on overall assembly defects (cobot-

320 related and human-related errors) were collected during the trials, as described in Section  
321 5. Additionally, data on the operators' stress response during the assembly phase were  
322 collected, as per Section 6.

323

#### 324 **4. COMPLEXITY ANALYSIS**

325 In scientific literature, complexity is typically used as a metric to predict  
326 production performance, including production times and defects. Indeed, it is often found  
327 that a reduction in complexity is associated with a significant performance improvement  
328 [41–43]. In this study, the structural complexity model, first introduced by Sinha et al. [44]  
329 and later adapted by Alkan and Harrison [45] and Verna et al. [43,46], serves as the  
330 foundation for assessing the assembly complexity of selected ARDUINO products. This  
331 model, originally developed for manual and fully automated assembly, is extended to the  
332 HRC assembly of the present case study, where the robot primarily performs  
333 organizational and logistical tasks, such as selecting components to be assembled in a  
334 predetermined sequence and delivering them to the human assembler. Adapting and  
335 integrating the structural complexity model to the domain of HRC assembly for highly  
336 customized product variants represents an innovative aspect of this study. This enables a  
337 quantitative assessment of assembly complexity within the context of mass  
338 customization.

339 The six product variants were selected to cover a wide range of assembly  
340 complexity. In the case study, each hole on the breadboard was modelled as a single  
341 component. This assumption allows multiple connections between the components and

342 the board to be modelled and distinguished from single connections. For example,  
 343 pushbuttons, i.e., the components that close a circuit when pressed, consist of four  
 344 different pins that need to be connected to the board. As this type of connection is more  
 345 complex than connecting a single-pin component, it was necessary to model the  
 346 individual holes on the board to distinguish these different cases.

347 The structural complexity model used to model the HRC assembly complexity is  
 348 based on Huckel's molecular theory [47] and defines the structural complexity of any  
 349 network-based engineering system as a function of the complexity of individual  
 350 components ( $C_1$ ), the pairwise interaction between connected components ( $C_2$ ), and the  
 351 effects of the overall system topology ( $C_3$ ). The structural complexity, represented as  $C$ ,  
 352 is a combination of these factors and can be expressed as:

$$C = C_1 + C_2 \cdot C_3. \quad (1)$$

353 In Eq. (1),  $C_1$  represents the complexity of managing and interacting with the  
 354 individual components of a product when they are considered separately, i.e., the  
 355 handling complexity of the product.  $C_1$  can be defined as follows:

$$C_1 = \sum_{p=1}^N h_p \quad (2)$$

356 where  $N$  is the total number of product components and  $h_p$  is the handling complexity of  
 357 component  $p$ . One of the most widely accepted models for calculating a handling  
 358 complexity index of individual components is the Lucas method [45], based on Design For  
 359 Assembly (DFA). This method uses a point scale that provides a relative measure of

360 assembly difficulty (a normalized handling complexity index) based on the size, weight,  
361 handling difficulty and orientation (alpha and beta symmetry) of individual components  
362 (see Table 3). Using the Lucas method, each component can be assigned a different  
363 handling complexity index (see Table 5). The higher the value of  $h_p$ , the more difficult the  
364 component is to handle and place on the board. These values are obtained as follows:

$$h_p = \frac{d_h^A + \sum_1^{N_B} d_h^B + d_h^C + d_h^D}{h_{max}} \quad (3)$$

365 where  $d_h^{i \in \{A,B,C,D\}}$  is the handling difficulty of attribute  $i$ ,  $N_B$  is the number of applicable  
366 handling difficulties related to attribute B, and  $h_{max}$  is the theoretical maximum value for  
367 the handling index (i.e., 6.9, according to Table 3).

368 In Eq. (1),  $C_2$  is the complexity of connections and liaisons between components,  
369 calculated as the sum of the complexities of the pairwise connections present in the  
370 product structure, according to Eq. (4):

$$C_2 = \sum_{p=1}^{N-1} \sum_{r=p+1}^N c_{pr} \cdot e_{pr} \quad (4)$$

371 where  $c_{pr}$  is the complexity in achieving a connection between components  $p$  and  $r$ , and  
372  $e_{pr}$  is the  $(p,r)^{th}$  entry of the binary adjacency matrix (**AM**) of the product. It has to be  
373 noted that in this specific case study, given that all components are connected to the  
374 breadboard, the  $r$ -th component is always the breadboard.

375 The complexity  $c_{pr}$  can be evaluated by the Lucas Method [45], by using the  
376 difficulty of connection attributes reported in Table 4, and is obtained as follows:

$$c_{pr} = \frac{d_c^E + d_c^F + d_c^G + d_c^H + d_c^I + d_c^J + d_c^K}{c_{\max}} \quad (5)$$

377 where  $d_c^{j \in \{E,F,G,H,I,J,K\}}$  is the connection difficulty of attribute  $j$ , and  $c_{\max}$  is the theoretical  
378 maximum value for the connection index (i.e., 13.1, according to Table 4).

379 Thus, the Lucas method provides a normalized assembly index that penalizes the  
380 physical attributes (e.g. component positioning and fastening, assembly direction,  
381 visibility, alignment and resistance to insertion) that affect assembly difficulty.

382 In Eq. (4),  $e_{pr}$  is defined by using the symmetric **AM** matrix of the product (see Fig.  
383 3). It can take two different values:

$$e_{pr} = \begin{cases} 1, & \text{if there is a connection between } p \text{ and } r \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

384 Each entry in the **AM** matrix indicates the presence of an assembly connection  
385 between the component and the breadboard. As an example, Fig. 3 shows the **AM** matrix  
386 of product variant V3.

387 As shown in Table 5, the connection complexity of each component with the  
388 breadboard ( $c_{pr}$ ) in the six electronic board variants (V1-V6) can take different values  
389 depending on multiple factors. For example, the connection complexity of long wires to  
390 the breadboard ranges from 3.7 to 6.3, depending on how the component is inserted into  
391 the breadboard and what other components are already connected. A complexity score  
392 of 5.3, for instance, is given if the wire needs to be bent to make the connection, and 6.3  
393 if the connection is made with reduced visibility.

394 Finally, in Eq. (1),  $C_3$  represents the topological complexity, i.e., the complexity  
395 associated with the product architecture pattern, which is defined as follows:

$$C_3 = \frac{E_{AM}}{N} = \frac{\sum_{q=1}^N \delta_q}{N}, \quad (7)$$

396 where  $E_{AM}$  is the matrix energy of  $\mathbf{AM}$ , i.e. the sum of the singular values  $\delta_q$  of  $\mathbf{AM}$  [43].  
397 It increases as the system topology shifts from centralized to more distributed  
398 architectures [44].

399 According to the increasing total assembly complexity  $C$ , Table 6 lists the  
400 complexities  $C_1$ ,  $C_2$  and  $C_3$  of the selected product variants. It is worth noting that an  
401 increase in complexity does not always imply an increase in the number of components.  
402 In fact, although variant V5 has more components than variant V6, the total complexity  
403 of variant V6 is higher than that of variant V5. This is due to the different nature of the  
404 components that compose the different products, the nature of the connections and the  
405 architecture of the final assembly.

406

## 407 5. PRODUCT AND PROCESS QUALITY ANALYSIS

408 In this section, complexity measures are integrated into the analysis of product  
409 and process quality, providing a novel perspective on the relationship between assembly  
410 complexity and the occurrence of defects in collaborative assembly processes for  
411 customized products.

412 During the manufacturing process, quality data on the overall defectiveness of  
413 product and process were collected to assess the quality of the HRC system (see Table 7).

414 Specifically, for each product variant assembly, the total number of defects (both in-  
415 process defects occurring during assembly – referred to as D1 - and offline defects  
416 detected during offline quality control – referred to as D2) was recorded. A classification  
417 was made for both types of defects, D1 and D2 (see Table 8). During the manufacturing  
418 process, the assembly operators and the quality control operator filled Table 8, indicating  
419 the number of defects found in each category for each assembled board. Certain defect  
420 categories, such as "Unpicked Component" and "Slipped Component," specifically relate  
421 to errors made by the cobot during the assembly phase. These categories reflect instances  
422 where the cobot failed to pick up a component or where a component slipped during the  
423 cobot handling. It is important to highlight that these defect categories capture cobot-  
424 related errors occurring during the assembly phase. Furthermore, it should be noted that  
425 the defects recorded in the in-process and offline phases reflect a combination of both  
426 cobot-related and human-related errors. This means that the defect data collected  
427 encompasses the performance of both the cobot and the human operators involved in  
428 the assembly process. To achieve a holistic view of the quality of the system, the total  
429 number of assembly defects  $D_{tot}$  (i.e., the sum of in-process and offline defects) were  
430 considered and analyzed (see Table 7).

431 The exclusion rule used was the Modified Interquartile Range Method, which is  
432 widely recognized as a practical and effective method for identifying outliers, taking into  
433 account the sample size [48]. The relationship between the total number of defects  
434 recorded by the six operators for each of the six variants of electronic boards and the  
435 complexity of the assembly (calculated as described in Section 4) was then analyzed. The

436 “operator factor” was not considered in the analysis after checking its non-significance at  
437 95% confidence level using a two-way ANOVA ( $p$ -value of 0.290). The Poisson regression  
438 model was used for the analysis, as total defects are count data [49]. The logarithm and  
439 square root link functions were considered, and different models were compared up to  
440 the third order of the predictor (i.e., assembly complexity  $C$ ). The selection of the best  
441 model was made based on Akaike's Corrected Information Criterion (AICc) and Bayesian  
442 Information Criterion (BIC), goodness-of-fit tests (Deviance and Pearson tests), and  
443 deviance residual plots [49,50]. The Deviance and Pearson tests assessed whether the  
444 predicted number of events deviated from the observed number in a way that was not  
445 predicted by the Poisson distribution. If the  $p$ -value was less than the significance level,  
446 the null hypothesis that the Poisson distribution provided a good fit could be rejected  
447 [49,50].

448 According to the results, the most appropriate Poisson model to describe the  
449 relationship between defects and complexity was the one using the square root link  
450 function, represented as:

$$D_{tot} = (k_1 \cdot C)^2, \quad (8)$$

451 where  $D_{tot}$  is the total number of defects (in-process and offline),  $C$  is assembly  
452 complexity evaluated according to Eq. (1), and  $k_1$  is the regression coefficient. The results  
453 of the Poisson regression analysis, reported in Table 9, showed that the relationship  
454 between  $D_{tot}$  and  $C$  was statistically significant. In addition, the analysis of the deviance  
455 residuals and the goodness-of-fit tests of Deviance and Pearson (where  $p$ -values are

456 higher than the significance level of 0.05) indicated that the model fitted the data well. In  
457 addition, a very high value of the deviance  $R^2$  was obtained.

458 Fig. 4(a) shows the total defects recorded during the experiment and the predicted  
459 curve obtained by Poisson regression with 95% confidence and prediction intervals are  
460 represented. Moreover, Fig. 4(b) shows the deviance residual plots, where the residuals  
461 appear satisfactory overall. Also using the Anderson-Darling test, the hypothesis of  
462 normality of the residual distribution cannot be rejected at the 95% confidence level ( $p$ -  
463 value = 0.194, which is higher than the significance level of 5%). The results obtained for  
464 product and process quality show that the increase in assembly complexity of the variants  
465 leads to an increase in the total number of defects, following a non-linear trend.

466

## 467 6. HUMAN WELL-BEING ANALYSIS

468 In this section, existing methodologies for assessing human well-being are  
469 integrated and adapted to capture the impact of assembly complexity on the human  
470 stress response in the context of mass customization, showcasing the originality of the  
471 proposed approach.

472 Physiological measures can be used to assess the state of human well-being during  
473 production, providing an objective measure compared to self-report tools, which may  
474 suffer from retrospective post-task bias [51]. Electrodermal activity (EDA) data is used in  
475 this study as a measure of human well-being, as it is commonly used as an indicator of  
476 the human stress response [52]. The Empatica E4 wristband (see Fig. 5(a)), a non-invasive  
477 biosensor that records EDA information at 4 Hz, was used to collect the EDA data. In

478 addition to EDA, the Empatica E4 also records information on pulsed blood volume (BVP),  
479 operator pulse motion (ACC), heart rate variability (HRV) and temperature (TMP). Fig. 5(b)  
480 shows an example of the raw output provided by the Empatica E4.

481 For each test performed by the operators, this raw signal was recorded and then  
482 analyzed using the EDA Explorer software [53]. This software removes any external noise  
483 from the raw signal and decomposes the EDA signal into two types of signals: the tonic  
484 signal and the phasic signal. The tonic signal refers to the long-term fluctuations of the  
485 EDA signal that are not explicitly triggered by external stimuli. Changes in Skin  
486 Conductance Level (SCL) are the best indicator of tonic activity. On the other hand, phasic  
487 activity refers to transient changes in EDA that are triggered by typically perceived and  
488 externally delivered stimuli. It is best characterized by Skin Conductance Response (SCR)  
489 changes. Accordingly, the SCR can be defined as a change in the amplitude of the EDA  
490 signal from the SCL to a peak in the response [52].

491 According to its widespread use [51,52], the average value of the SCR peak  
492 amplitude was used as a stress indicator for each assembly worker in this study. The peak  
493 amplitude values were then normalized in the formulation of the final stress indicator to  
494 remove individual differences between individuals. As a result, the human stress response  
495 ( $H_S$ ) indicator for each operator can be defined as:

$$H_S = \left[ \frac{\frac{\sum_{w=1}^{N_P} a_w}{N_P} - a_{min}}{a_{max} - a_{min}} \right] \cdot 100, \quad (9)$$

496 where  $a_w$  is the amplitude of the  $w$ -th SCR peak,  $N_P$  is the total number of SCR peaks  
497 during the assembly of a given product variant,  $a_{min}$  is the minimum amplitude of the

498 SRC peaks and  $a_{max}$  is the maximum amplitude of the SRC peaks (both related to each  
499 operator).

500 The human stress response data obtained during the 36 assembly processes (i.e.,  
501 the 6 product variant assemblies performed by the 6 operators) are reported in Table 7.  
502 The  $H_S$  value of each operator is related to the assembly complexity (as per Section 4) in  
503 order to model the function that captures their relationship. The “operator factor” was  
504 not considered in the analysis after checking its non-significance at 95% confidence level  
505 using a two-way ANOVA ( $p$ -value of 0.999). Fig. 6 shows the two-term power curve fitting  
506 relating human stress response and product variant assembly complexity, in the form:

$$H_S = k_2 \cdot C^{k_3}, \quad (10)$$

507 where  $H_S$  is the human stress response,  $C$  is the assembly complexity evaluated according  
508 to Eq. (1), and  $k_2$  and  $k_3$  are the regression coefficients.

509 This model was the best-fitting model compared to various linear and non-linear  
510 models, considering the goodness-of-fit statistics and residual analysis [54]. The statistical  
511 significance of the parameter estimate is confirmed by checking that the 95% confidence  
512 intervals for the parameters, calculated from the corresponding Standard Errors (SE)  
513 reported in Table 10, exclude the zero [55,56]. The  $S$ -value, i.e., the standard error of the  
514 regression, is a measure of the goodness of fit of the model under consideration instead  
515 of the  $R^2$  for non-linear models [56]. The residual plots in Fig. 6(b) appear satisfactory  
516 overall and, using the Anderson-Darling test, the hypothesis of normality of the residual  
517 distribution cannot be rejected at the 95% confidence level. It should be noted that,  
518 according to the result obtained, non-linear regression is preferable to linear quadratic

519 regression, as linearizing the function to perform linear regression can lead to bias in the  
520 predictions [57]. According to the results shown in Table 10 and Fig. 6, there is a super-  
521 linear relationship between human stress response and the complexity of product variant  
522 assembly. This result, which is one of the first attempts to study the relationship between  
523 assembly complexity and human stress response, shows that as the complexity of the  
524 product assembly increases, the assembly process becomes more challenging and entails  
525 a higher degree of mental workload and cognitive effort, leading to a more than  
526 proportional increase in human stress response.

527

## 528 **7. HRC-QWAT**

529 This section introduces the "Human-Robot Collaboration Quality and Well-Being  
530 Assessment Tool" (HRC-QWAT), a tool designed to synthesize previous analyses of quality  
531 and human well-being, by directly relating  $H_S$  and  $D_{tot}$ , regardless of the complexity of  
532 the product assembled. The selection of total defects and human stress response in the  
533 HRC-QWAT tool was based on their significant impact on evaluating the performance of  
534 the HRC system, including both product quality and human well-being. The tool  
535 establishes a direct relationship between human stress response and total defects,  
536 enabling a comprehensive assessment of system performance within the context of  
537 single-variant and small-batch variant production. This design is specifically tailored to  
538 address the distinct challenges and demands posed by customized production scenarios,  
539 where the adaptability of the production process and the individuality of each assembly  
540 play pivotal roles.

541 Two typologies of HRC-QWAT are proposed. The first typology is intended for  
542 single variant production of highly customized products produced one by one in the HRC  
543 system, even if repeated over time. This type of production involves the manufacture of  
544 a single product variant at a time, typically in response to specific customer orders or  
545 market demand. The production process is adapted as required for each variant, which  
546 can result in longer lead times and higher production costs. In this scenario, the company  
547 is interested in monitoring the performance of each individual product variant assembly  
548 in terms of quality and human well-being. On the other hand, the second typology of HRC-  
549 QWAT is proposed to provide companies with a diagnostic method for products of the  
550 same variant manufactured in small batches, after each of such productions. This type of  
551 production involves the manufacture of small batches of the same product variant,  
552 typically in response to forecasted demand or market trends. The production process is  
553 adapted for each batch, allowing a product variant to be produced more efficiently and  
554 cost-effectively than in the single variant scenario. The choice between single variant  
555 production and small-batch variant production generally depends on factors such as  
556 demand variability, lead time requirements, and production costs. Single variant  
557 production is best suited for highly customized products with low demand, while variant  
558 batch production is more efficient for producing a range of products with moderate to  
559 high demand.

560 The use phase of the proposed tool is the same for practitioners in both cases,  
561 while the difference lies in the realization phase of the HRC-QWAT.

562 In both typologies, the HRC-QWAT allows for the assessment of quality and human  
563 well-being in the collaborative assembly process, taking into account the unique  
564 characteristics and requirements of each production scenario. This tool offers a  
565 comprehensive evaluation of the HRC system's performance, considering the relationship  
566 between human stress response ( $H_S$ ), total defects ( $D_{tot}$ ), and the complexity of the  
567 assembly process. Although the complexity indicator is not explicitly included as a  
568 separate metric in the HRC-QWAT, its influence on the performance measures is implicitly  
569 accounted for in the evaluation. As discussed in the previous sections, assembly  
570 complexity plays a crucial role in affecting performance metrics, particularly in highly  
571 customized and personalized product assemblies within the same product family [58–60].  
572 Although the HRC-QWAT does not directly measure complexity, it considers its impact on  
573 the overall performance of the collaborative process. By capturing the relationship  
574 between human stress response, total defects, and the intricate nature of the assembly,  
575 the tool indirectly accounts for the effects of assembly complexity on the HRC system's  
576 performance.

577 To construct the HRC-QWAT, the following operational steps should be taken.  
578 First, a set of historical experimental data representative of the production must be  
579 collected. In the case of the HRC-QWAT for single variant production, a reasonable  
580 number of products (at least about thirty, for robust regression parameter estimation  
581 [61]) should be produced, and quality and human stress responses should be measured  
582 (according to Section 5 and 6, respectively). On the other hand, for the HRC-QWAT for  
583 small batches, an adequate number of production units should be collected for each

584 batch (at least about fifteen units for each product type, if possible [61]) and the average  
585 performance measures should be obtained for each batch. As mentioned in Section 5 and  
586 6, it is advisable to perform preliminary data analysis using conventional statistical  
587 techniques to detect and filter outliers [62].

588 Second, the model relating the two performance measures should be developed  
589 to represent the overall quality of the systems, in terms of product/process quality and  
590 human well-being. Considering the case study, the combination of the models in Eq. (8)  
591 and (10) leads to a linear model. Such a linear model is the best fit when considering single  
592 variant production, as also confirmed by the goodness-of-fit statistics and residual  
593 analysis [54]. Fig. 7(a) depicts the prediction model relating human stress response  $H_S$  to  
594 total defects  $D_{tot}$  and Fig. 7(b) shows the residual plots. The output of the regression is  
595 shown in Table 11.

596 When considering small batches of products of the same variant, average values  
597 of human stress response ( $\overline{H_S}$ ) and total defects ( $\overline{D_{tot}}$ ) should be obtained for each  
598 variant. Then, the prediction model should be derived using these averages. In the case  
599 study, six small batches are considered, one for each product variant (V1-V6), each  
600 consisting of six products.

601 Referring to the case study data, Fig. 8 illustrates the best fitting model, i.e., a  
602 linear regression model, with the residual plots, and the main output of the regression is  
603 reported in Table 11.

604 The HRC-QWAT diagnostic tool (see Fig. 9) uses the model as a reference for  
605 prediction and takes into account the associated uncertainty range. Specifically, the two

606 prediction limits (Lower Prediction Limit *LPL* and Upper Prediction Limit *UPL*) derived from  
 607 the regression models shown in Fig. 7 and 8, serve as thresholds for identifying critical  
 608 products and small batches, respectively. Products and small batches are classified as  
 609 critical in terms of both defects and human stress response when a special source of  
 610 variation i.e., a source not inherent to the process, occurs [61]. It should be noted that  
 611 the negative values of *LPL* are set equal to zero, as this is not physically possible. As a  
 612 result, for some products or batches, the prediction interval may not be symmetrical with  
 613 respect to the predicted regression value, as shown in Fig. 9.

614 The two prediction limits can be calculated as follows:

$$LPL = \widehat{H}_S - t_{1-\frac{\alpha}{2}, \gamma} \sqrt{[SE(Fit)]^2 + S^2} \quad (11)$$

$$UPL = \widehat{H}_S + t_{1-\frac{\alpha}{2}, \gamma} \sqrt{[SE(Fit)]^2 + S^2}$$

615 where  $\widehat{H}_S$  is the predicted value of the regression curve,  $t_{1-\frac{\alpha}{2}, \gamma}$  is the value of the Student's  
 616 *t* distribution with  $\gamma$  degrees of freedom (i.e., number of observations minus 1) and  
 617 significance level  $\alpha$ ,  $SE(Fit)$  is the standard error of the fit, and  $S$  is the standard error of  
 618 the regression [61].

619 In the use phase, when new single products or small batches of products are  
 620 produced, the observed values  $(D_{tot}, H_S)$  or  $(\overline{D}_{tot}, \overline{H}_S)$  are compared with the  
 621 corresponding prediction limits from the HRC-QWAT for single variant or small-batch  
 622 production, respectively. Accordingly:

623 a) If the observed  $(D_{tot}, H_S)$  or  $(\overline{D}_{tot}, \overline{H}_S)$  value falls within the prediction  
 624 range (*LPL*, *UPL*), the product or batch is considered non-critical.

625           b) If the observed  $(D_{tot}, H_S)$  or  $(\overline{D}_{tot}, \overline{H}_S)$  value is higher than the upper  
626 prediction limit (*UPL*) (area A in Fig. 9) or lower than the lower prediction limit (*LPL*)  
627 (area B in Fig. 9), it indicates a mismatch between the human stress response and the  
628 total defects, and an abnormal situation exists, resulting in the product or batch being  
629 signaled as critical. Specifically, products or batches located in area A of Fig. 9 are  
630 reported as critical due to the high level of stress response experienced by operators  
631 compared to the number of total defects detected. On the other hand, products or  
632 batches lying in area B are characterized by abnormal defectiveness compared to the  
633 level of human stress response.

634           Table 12 reports an example of critical product production and an example of  
635 small-batch production detected as critical using the HRC-QWAT and possible root causes.

636           The proposed diagnostic tool has been developed with a dual objective. Firstly, it  
637 aims to accurately position products or small batches on the HRC-QWAT, thereby  
638 providing a clear understanding of their relative position compared to other products.  
639 This information can be valuable in making informed quality control decisions and  
640 identifying areas for improvement. Secondly, the diagnostic tool is designed to detect  
641 unusual production scenarios and identify critical out-of-control situations. By  
642 continuously monitoring production processes, the tool can identify any deviations from  
643 established normal operating conditions, allowing corrective action to be taken in a timely  
644 manner. This feature of the diagnostic tool acts as an in-process control mechanism,  
645 ensuring that the quality of the overall system (product/process and human) remains  
646 consistently high throughout the production process.

647 In conclusion, the proposed diagnostic tool represents a significant step forward  
648 in quality control and monitoring, providing manufacturers with a powerful tool to ensure  
649 consistent product quality and to detect and correct quality deviations in real-time.

650

## 651 **8. DISCUSSION**

652 The novelty of the HRC-QWAT lies in its comprehensive assessment of quality  
653 systems, encompassing both technical aspects of production quality and the human  
654 factor of worker well-being. While previous studies have focused on individual  
655 measurements and indicators in HRC systems, the HRC-QWAT combines multiple  
656 dimensions of evaluation to provide a more holistic understanding of the collaborative  
657 assembly process. By integrating indicators related to total defects and human stress  
658 response, the tool offers a more nuanced evaluation of the performance of HRC systems.

659 Moreover, the HRC-QWAT's versatility and adaptability contribute to its novelty.  
660 It can be applied to both single variant and small-batch production scenarios,  
661 accommodating different production environments and collaboration settings. Whether  
662 the work is predominantly performed by a robot or in a high-intensity human work  
663 environment, the HRC-QWAT assesses production quality and worker stress response,  
664 ensuring an optimized collaborative process. The tool's adaptability allows it to be fine-  
665 tuned to the unique parameters of various production environments and collaboration  
666 scenarios, making it not only a quality and well-being assessment tool but also a strategic  
667 tool for comparing and contrasting different collaboration scenarios.

668           The potential for generalization is another key aspect of the HRC-QWAT's novelty.  
669 Although the case study focused on electronic board assembly, the design and  
670 methodology of the HRC-QWAT were conceived with a broader application in mind. Its  
671 adaptability allows it to be utilized in a wide range of production scenarios, even when a  
672 robot performs the majority of the work and the role of the human operator is minimal  
673 or focused on labor-intensive tasks. It should be noted, however, that the generalizability  
674 of the HRC-QWAT depends on careful adaptation and refinement of the model  
675 parameters. This will allow the tool to accurately reflect the interaction dynamics and  
676 associated stress responses in different HRC settings. The possibility of extending the use  
677 of the HRC-QWAT to more diverse and nuanced collaboration scenarios represent a  
678 promising avenue for future research and development in the field of HRC.

679

## 680 **9. CONCLUSIONS**

681           The aim of the present research was to propose a novel tool, called the HRC-  
682 QWAT, which combines two indicators to evaluate and monitor the quality of a  
683 production system: the total number of defects generated during the production of  
684 product variants, and the stress response of workers. This innovative tool addresses a  
685 significant gap in the field of human-robot collaboration assessment, providing a unique  
686 approach to evaluating both the production quality and the well-being of human  
687 operators. The methodology used a collaborative human-robot assembly system as a case  
688 study to demonstrate the feasibility of the HRC-QWAT approach. The methodology  
689 consists of two main phases: (1) the realization phase, in which the HRC-QWAT is

690 constructed by collecting historical experimental data and developing a model relating  
691 the two performance measures (total defects and human stress response) that represent  
692 the overall quality of the system; and (2) the use phase, in which the HRC-QWAT is used  
693 as a reference for predicting future products/batches and identifying critical products in  
694 terms of defects and human stress response. The diagnostic tool uses the model to  
695 compare observed performance measures with corresponding prediction limits and  
696 detect abnormal production scenarios.

697 The HRC-QWAT introduces a novel approach to the evaluation of quality systems  
698 in HRC. Unlike previous studies that focused on individual metrics, this tool  
699 comprehensively assesses both technical production quality and worker well-being  
700 factors. Its adaptability and versatility make it suitable for single variant or small-batch  
701 production, and for different environments and collaborative settings. In addition,  
702 although in this study the HRC-QWAT was applied to electronic boards assembly, its  
703 adaptable design allows for a broader application, opening doors for future research in  
704 the evaluation and development of human-robot collaboration.

705 A limitation of the proposed approach is the use of a structural complexity model,  
706 which was originally designed for manual and fully automated processes. While such a  
707 model serves as a good first approximation for collaborative human-robot contexts, as  
708 the cobot mainly performs logistical and organizational support tasks, a more refined  
709 complexity model will be required for a more accurate evaluation. Another limitation of  
710 the study is that the comparison in the HRC-QWAT is based on only two indicators, total  
711 defects and human stress response. In such measures, the performance of the cobot is

712 not directly evaluated, although it is implicitly reflected in the total number of defects.  
713 Furthermore, additional performance measures, such as workload, are not directly  
714 addressed. Recognizing these limitations, it is important to consider that the proposed  
715 approach has the flexibility to be extended to include additional indicators, including  
716 cobot performance and workload, as well as encompass process sustainability and  
717 economic impact measures.

718 Future research efforts will aim to overcome (at least some of) the above  
719 limitations. Particular attention will be paid to refining the complexity model by including  
720 factors related to HRC and performing a validation of the proposed approach using  
721 different products to quantitatively assess its efficiency. In addition, the study could be  
722 extended to include other cobot performance measures, including efficiency metrics  
723 (cycle time, throughput), accuracy and reliability metrics, safety metrics, and  
724 environmental/economic sustainability indicators, such as equivalent carbon dioxide  
725 emissions and life cycle costs.

726

## 727 NOMENCLATURE

728

HRC	Human-Robot Collaboration
HRC-QWAT	Human-Robot Collaboration Quality and Well-Being Assessment
	Tool
$C$	Assembly complexity
$C_1$	Handling complexity

$N$	Total number of product components
$h_p$	Handling complexity of component $p$
$d_h^{i \in \{A,B,C,D\}}$	Handling difficulty of attribute $i$
$h_{\max}$	Theoretical maximum value for the handling index
$N_B$	Number of applicable handling difficulties related to attribute B
$C_2$	Complexity of connections and liaisons
$c_{pr}$	Complexity in achieving a connection between components $p$ and $r$
<b>AM</b>	Binary adjacency matrix of the product
$e_{pr}$	$(p,r)^{\text{th}}$ entry of the <b>AM</b> matrix of the product
$d_c^{j \in \{E,F,G,H,I,J,K\}}$	Connection difficulty of attribute $j$
$c_{\max}$	Theoretical maximum value for the connection index
$e_{pr}$	$e_{pr} = \begin{cases} 1, & \text{if there is a connection between } p \text{ and } r \\ 0, & \text{otherwise} \end{cases}$
$C_3$	Topological complexity
$E_{AM}$	Matrix energy of <b>AM</b>
$\delta_q$	Singular values of <b>AM</b>
D1	In-process defects
D2	Offline defects
$D_{tot}$	Total number of defects

$k_1$	Regression coefficient of the model $D_{tot}$ vs $C$
EDA	Electrodermal activity
SCR	Skin Conductance Response
$H_S$	Human stress response
$a_w$	Amplitude of the $w$ -th SCR peak
$N_P$	Total number of SCR peaks
$a_{min}$	Minimum amplitude of the SRC peaks
$a_{max}$	Maximum amplitude of the SRC peaks
$k_2$	Regression coefficient of the model $H_S$ vs $C$
$k_3$	Regression coefficient of the model $H_S$ vs $C$
CI	Confidence Interval
SE	Standard Error
$R^2$	Coefficient of determination
$S$	Standard error of the regression
$\overline{H_S}$	Average value of human stress response
$\overline{D_{tot}}$	Average value of total defects
$k_4$	Regression coefficient of the model $H_S$ vs $D_{tot}$
PI	Prediction Interval
LPL	Lower Prediction Limit and

$UPL$	Upper Prediction Limit
$\widehat{H}_S$	Predicted $H_S$ value of the regression curve
$t_{1-\frac{\alpha}{2}, \gamma}$	Value of the Student's $t$ distribution with $\gamma$ degrees of freedom and significance level $\alpha$
$SE(Fit)$	Standard error of the fit

729

730 **REFERENCES**

- 731 [1] Falck, A.-C., Örtengren, R., Rosenqvist, M., and Söderberg, R., 2017, "Basic  
732 Complexity Criteria and Their Impact on Manual Assembly Quality in Actual  
733 Production," *Int. J. Ind. Ergon.*, **58**, pp. 117–128.
- 734 [2] Buckholtz, B., Ragai, I., and Wang, L., 2015, "Cloud Manufacturing: Current Trends  
735 and Future Implementations," *J. Manuf. Sci. Eng.*, **137**(4).
- 736 [3] Krüger, J., Lien, T. K., and Verl, A., 2009, "Cooperation of Human and Machines in  
737 Assembly Lines," *CIRP Ann.*, **58**(2), pp. 628–646.
- 738 [4] Peshkin, M., and Colgate, J. E., 1999, "Cobots," *Ind. Robot An Int. J.*, **26**(5), pp.  
739 335–341.
- 740 [5] Maddikunta, P. K. R., Pham, Q.-V., Prabadevi, B., Deepa, N., Dev, K., Gadekallu, T.  
741 R., Ruby, R., and Liyanage, M., 2021, "Industry 5.0: A Survey on Enabling  
742 Technologies and Potential Applications," *J. Ind. Inf. Integr.*, p. 100257.
- 743 [6] Coronado, E., Kiyokawa, T., Ricardez, G. A. G., Ramirez-Alpizar, I. G., Venture, G.,  
744 and Yamanobe, N., 2022, "Evaluating Quality in Human-Robot Interaction: A  
745 Systematic Search and Classification of Performance and Human-Centered  
746 Factors, Measures and Metrics towards an Industry 5.0," *J. Manuf. Syst.*, **63**, pp.  
747 392–410.
- 748 [7] Ramanujan, D., Bernstein, W. Z., Diaz-Elsayed, N., and Haapala, K. R., 2023, "The  
749 Role of Industry 4.0 Technologies in Manufacturing Sustainability Assessment," *J.*  
750 *Manuf. Sci. Eng.*, **145**(1).
- 751 [8] Deissenboeck, F., Juergens, E., Lochmann, K., and Wagner, S., 2009, "Software  
752 Quality Models: Purposes, Usage Scenarios and Requirements," *2009 ICSE*  
753 *Workshop on Software Quality*, IEEE, pp. 9–14.

- 754 [9] Damacharla, P., Javaid, A. Y., Gallimore, J. J., and Devabhaktuni, V. K., 2018,  
755 “Common Metrics to Benchmark Human-Machine Teams (HMT): A Review,” IEEE  
756 Access, **6**, pp. 38637–38655.
- 757 [10] Breque, M., De Nul, L., and Petridis, A., 2021, “Industry 5.0: Towards a  
758 Sustainable, Human-Centric and Resilient European Industry,” Luxemb. LU Eur.  
759 Comm. Dir. Res. Innov.
- 760 [11] Leng, J., Sha, W., Wang, B., Zheng, P., Zhuang, C., Liu, Q., Wuest, T., Mourtzis, D.,  
761 and Wang, L., 2022, “Industry 5.0: Prospect and Retrospect,” J. Manuf. Syst., **65**,  
762 pp. 279–295.
- 763 [12] Xu, X., Lu, Y., Vogel-Heuser, B., and Wang, L., 2021, “Industry 4.0 and Industry  
764 5.0—Inception, Conception and Perception,” J. Manuf. Syst., **61**, pp. 530–535.
- 765 [13] Marvel, J. A., Bagchi, S., Zimmerman, M., and Antonishek, B., 2020, “Towards  
766 Effective Interface Designs for Collaborative HRI in Manufacturing: Metrics and  
767 Measures,” ACM Trans. Human-Robot Interact., **9**(4), pp. 1–55.
- 768 [14] Khavas, Z. R., Ahmadzadeh, S. R., and Robinette, P., 2020, “Modeling Trust in  
769 Human-Robot Interaction: A Survey,” *Social Robotics: 12th International  
770 Conference, ICSR 2020, Golden, CO, USA, November 14–18, 2020, Proceedings 12*,  
771 Springer, pp. 529–541.
- 772 [15] Venkatesh, V., and Davis, F. D., 2000, “A Theoretical Extension of the Technology  
773 Acceptance Model: Four Longitudinal Field Studies,” Manage. Sci., **46**(2), pp. 186–  
774 204.
- 775 [16] Young, M. S., Brookhuis, K. A., Wickens, C. D., and Hancock, P. A., 2015, “State of  
776 Science: Mental Workload in Ergonomics,”  
777 <https://doi.org/10.1080/00140139.2014.956151>, **58**(1), pp. 1–17.
- 778 [17] Heard, J., Harriott, C. E., and Adams, J. A., 2018, “A Survey of Workload  
779 Assessment Algorithms,” IEEE Trans. Human-Machine Syst., **48**(5), pp. 434–451.
- 780 [18] Tabrez, A., Luebbbers, M. B., and Hayes, B., 2020, “A Survey of Mental Modeling  
781 Techniques in Human–Robot Teaming,” Curr. Robot. Reports, **1**, pp. 259–267.
- 782 [19] Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., and Cannon-Bowers, J. A.,  
783 2000, “The Influence of Shared Mental Models on Team Process and  
784 Performance.,” J. Appl. Psychol., **85**(2), p. 273.
- 785 [20] Hudlicka, E., 2003, “To Feel or Not to Feel: The Role of Affect in Human–  
786 Computer Interaction,” Int. J. Hum. Comput. Stud., **59**(1–2), pp. 1–32.
- 787 [21] Zhang, P., 2013, “The Affective Response Model: A Theoretical Framework of

- 788 Affective Concepts and Their Relationships in the ICT Context,” *MIS Q.*, pp. 247–  
789 274.
- 790 [22] Naneva, S., Sarda Gou, M., Webb, T. L., and Prescott, T. J., 2020, “A Systematic  
791 Review of Attitudes, Anxiety, Acceptance, and Trust towards Social Robots,” *Int. J.*  
792 *Soc. Robot.*, **12**(6), pp. 1179–1201.
- 793 [23] Lorenzini, M., Kim, W., and Ajoudani, A., 2022, “An Online Multi-Index Approach  
794 to Human Ergonomics Assessment in the Workplace,” *IEEE Trans. Human-*  
795 *Machine Syst.*, **52**(5), pp. 812–823.
- 796 [24] Ajoudani, A., Albrecht, P., Bianchi, M., Cherubini, A., Del Ferraro, S., Fraise, P.,  
797 Fritzsche, L., Garabini, M., Ranavolo, A., and Rosen, P. H., 2020, “Smart  
798 Collaborative Systems for Enabling Flexible and Ergonomic Work Practices  
799 [Industry Activities],” *IEEE Robot. Autom. Mag.*, **27**(2), pp. 169–176.
- 800 [25] Fan, J., Zheng, P., and Lee, C. K. M., 2023, “A Vision-Based Human Digital Twin  
801 Modelling Approach for Adaptive Human-Robot Collaboration,” *J. Manuf. Sci.*  
802 *Eng.*, pp. 1–11.
- 803 [26] Verna, E., Puttero, S., Genta, G., and Galetto, M., 2023, “Toward a Concept of  
804 Digital Twin for Monitoring Assembly and Disassembly Processes,” *Qual. Eng.*, pp.  
805 1–18.
- 806 [27] Irfan, B., Ramachandran, A., Spaulding, S., Glas, D. F., Leite, I., and Koay, K. L.,  
807 2019, “Personalization in Long-Term Human-Robot Interaction,” *2019 14th*  
808 *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, IEEE, pp.  
809 685–686.
- 810 [28] Müller, J., 2020, “Enabling Technologies for Industry 5.0—Results of a Workshop  
811 with Europe’s Technology Leaders,” *Dir. Res. Innov.*
- 812 [29] Hu, Y., Abe, N., Benallegue, M., Yamanobe, N., Venture, G., and Yoshida, E., 2022,  
813 “Toward Active Physical Human–Robot Interaction: Quantifying the Human State  
814 during Interactions,” *IEEE Trans. Human-Machine Syst.*, **52**(3), pp. 367–378.
- 815 [30] Setchi, R., Dehkordi, M. B., and Khan, J. S., 2020, “Explainable Robotics in Human-  
816 Robot Interactions,” *Procedia Comput. Sci.*, **176**, pp. 3057–3066.
- 817 [31] Anjomshoae, S., Najjar, A., Calvaresi, D., and Främpling, K., 2019, “Explainable  
818 Agents and Robots: Results from a Systematic Literature Review,” *18th*  
819 *International Conference on Autonomous Agents and Multiagent Systems*  
820 *(AAMAS 2019), Montreal, Canada, May 13–17, 2019*, International Foundation for  
821 Autonomous Agents and Multiagent Systems, pp. 1078–1088.
- 822 [32] Hoffman, G., 2019, “Evaluating Fluency in Human–Robot Collaboration,” *IEEE*

- 823 Trans. Human-Machine Syst., **49**(3), pp. 209–218.
- 824 [33] Heard, J., Harriott, C. E., and Adams, J. A., 2017, “A Human Workload Assessment  
825 Algorithm for Collaborative Human-Machine Teams,” *2017 26th IEEE*  
826 *International Symposium on Robot and Human Interactive Communication (RO-*  
827 *MAN)*, IEEE, pp. 366–371.
- 828 [34] Heard, J., Heald, R., Harriott, C. E., and Adams, J. A., 2019, “A Diagnostic Human  
829 Workload Assessment Algorithm for Collaborative and Supervisory Human--Robot  
830 Teams,” *ACM Trans. Human-Robot Interact.*, **8**(2), pp. 1–30.
- 831 [35] Petersen, S. A., Mannhardt, F., Oliveira, M., and Torvatn, H., 2018, “A Framework  
832 to Navigate the Privacy Trade-Offs for Human-Centred Manufacturing,”  
833 *Collaborative Networks of Cognitive Systems: 19th IFIP WG 5.5 Working*  
834 *Conference on Virtual Enterprises, PRO-VE 2018, Cardiff, UK, September 17-19,*  
835 *2018, Proceedings 19*, Springer, pp. 85–97.
- 836 [36] Mannhardt, F., Petersen, S. A., and Oliveira, M. F., 2019, “A Trust and Privacy  
837 Framework for Smart Manufacturing Environments,” *J. Ambient Intell. Smart*  
838 *Environ.*, **11**(3), pp. 201–219.
- 839 [37] Rahman, S. M. M., 2021, “Cybersecurity Metrics for Human-Robot Collaborative  
840 Automotive Manufacturing,” *2021 IEEE International Workshop on Metrology for*  
841 *Automotive (MetroAutomotive)*, IEEE, pp. 254–259.
- 842 [38] Causo, A., Durham, J., Hauser, K., Okada, K., and Rodriguez, A., 2020, *Advances on*  
843 *Robotic Item Picking*, Springer.
- 844 [39] Fujita, M., Domae, Y., Noda, A., Garcia Ricardez, G. A., Nagatani, T., Zeng, A.,  
845 Song, S., Rodriguez, A., Causo, A., Chen, I. M., and Ogasawara, T., 2019, “What  
846 Are the Important Technologies for Bin Picking? Technology Analysis of Robots in  
847 Competitions Based on a Set of Performance Metrics,”  
848 <https://doi.org/10.1080/01691864.2019.1698463>, **34**(7–8), pp. 560–574.
- 849 [40] Zadeh, L., 1962, “From Circuit Theory to System Theory,” *Proc. IRE*, **50**(5), pp.  
850 856–865.
- 851 [41] ElMaraghy, H., Schuh, G., ElMaraghy, W., Piller, F., Schönsleben, P., Tseng, M.,  
852 and Bernard, A., 2013, “Product Variety Management,” *Cirp Ann.*, **62**(2), pp. 629–  
853 652.
- 854 [42] Genta, G., Galetto, M., and Franceschini, F., 2018, “Product Complexity and  
855 Design of Inspection Strategies for Assembly Manufacturing Processes,” *Int. J.*  
856 *Prod. Res.*, **56**(11), pp. 4056–4066.
- 857 [43] Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022, “Defect Prediction

- 858 for Assembled Products: A Novel Model Based on the Structural Complexity  
859 Paradigm,” *Int. J. Adv. Manuf. Technol.*, **120**(5–6), pp. 3405–3426.
- 860 [44] Sinha, K., 2014, “Structural Complexity and Its Implications for Design of Cyber-  
861 Physical Systems,” PhD dissertation, Engineering Systems Division, Massachusetts  
862 Institute of Technology.
- 863 [45] Alkan, B., and Harrison, R., 2019, “A Virtual Engineering Based Approach to Verify  
864 Structural Complexity of Component-Based Automation Systems in Early Design  
865 Phase,” *J. Manuf. Syst.*, **53**, pp. 18–31.
- 866 [46] Verna, E., Genta, G., Galetto, M., and Franceschini, F., 2022, “Defects-per-Unit  
867 Control Chart for Assembled Products Based on Defect Prediction Models,” *Int. J.*  
868 *Adv. Manuf. Technol.*, **119**(5–6), pp. 2835–2846.
- 869 [47] Hückel, E., 1932, “Quantentheoretische Beiträge Zum Problem Der Aromatischen  
870 Und Ungesättigten Verbindungen. III,” *Zeitschrift für Phys.*, **76**(9–10), pp. 628–  
871 648.
- 872 [48] Barbato, G., Barini, E. M., Genta, G., and Levi, R., 2011, “Features and  
873 Performance of Some Outlier Detection Methods,”  
874 <http://dx.doi.org/10.1080/02664763.2010.545119>, **38**(10), pp. 2133–2149.
- 875 [49] Cameron, A. C., and Trivedi, P. K., 2013, *Regression Analysis of Count Data*,  
876 Cambridge university press.
- 877 [50] Myers, R. H., Montgomery, D. C., Vining, G. G., and Robinson, T. J., 2012,  
878 *Generalized Linear Models: With Applications in Engineering and the Sciences*,  
879 John Wiley & Sons, Hoboken, NJ, USA.
- 880 [51] Gervasi, R., Aliev, K., Mastrogiacomo, L., and Franceschini, F., 2022, “User  
881 Experience and Physiological Response in Human-Robot Collaboration: A  
882 Preliminary Investigation,” *J. Intell. Robot. Syst.*, **106**(2), p. 36.
- 883 [52] Zhao, B., Wang, Z., Yu, Z., and Guo, B., 2018, “EmotionSense: Emotion  
884 Recognition Based on Wearable Wristband,” *2018 IEEE SmartWorld, Ubiquitous  
885 Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing &  
886 Communications, Cloud & Big Data Computing, Internet of People and Smart City  
887 Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, IEEE, pp. 346–  
888 355.
- 889 [53] Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., and Picard, R., 2015,  
890 “Automatic Identification of Artifacts in Electrodermal Activity Data,” *Proc. Annu.*  
891 *Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, **2015-Novem**, pp. 1934–1937.
- 892 [54] Montgomery, D., Runger, G., and Hubele, N., 2010, *Engineering Statistics*, John

- 893 Wiley & Sons Inc., New York.
- 894 [55] Seber, G. A. F., and Wild, C. J., 1989, *Nonlinear Regression*, John Wiley & Sons,  
895 New York.
- 896 [56] Bates, D. M., and Watts, D. G., eds., 1988, *Nonlinear Regression Analysis and Its*  
897 *Applications*, John Wiley & Sons, Inc., Hoboken, NJ, USA.
- 898 [57] Galetto, M., Verna, E., and Genta, G., 2020, "Accurate Estimation of Prediction  
899 Models for Operator-Induced Defects in Assembly Manufacturing Processes,"  
900 *Qual. Eng.*, **32**(4), pp. 595–613.
- 901 [58] Hasan, S. M., Baqai, A. A., Butt, S. U., and quz Zaman, U. K., 2018, "Product Family  
902 Formation Based on Complexity for Assembly Systems," *Int. J. Adv. Manuf.*  
903 *Technol.*, **95**(1), pp. 569–585.
- 904 [59] Lim, K. Y. H., Zheng, P., Chen, C. H., and Huang, L., 2020, "A Digital Twin-Enhanced  
905 System for Engineering Product Family Design and Optimization," *J. Manuf. Syst.*,  
906 **57**, pp. 82–93.
- 907 [60] Dan, B., and Tseng, M. M., 2007, "Assessing the Inherent Flexibility of Product  
908 Families for Meeting Customisation Requirements," *Int. J. Manuf. Technol.*  
909 *Manag.*, **10**(2–3), pp. 227–246.
- 910 [61] Montgomery, D. C., 2019, *Introduction to Statistical Quality Control*, Wiley Global  
911 Education.
- 912 [62] Barbato, G., Germak, A., and Genta, G., 2013, *Measurements for Decision Making*,  
913 Società Editrice Esculapio, Bologna.
- 914 [63] Chan, V., and Salustri, F. A., 2003, "Dfa: The Lucas Method," Ryerson University,  
915 Toronto.
- 916

917

## Figure Captions List

918

- Fig. 1 Example of an assembled electronic board (variant V3): (a) final product assembled and (b) circuit diagram
- Fig. 2 Collaborative assembly workstation showing (a) the single-armed UR3e cobot equipped with the OnRobot RG6 gripper and (b) product components
- Fig. 3 **AM** matrix of variant V3
- Fig. 4 Total defects ( $D_{tot}$ ) vs assembly complexity ( $C$ ): (a) Poisson regression model and (b) Deviance residual plots
- Fig. 5 (a) Empatica E4 wristband (b) Empatica E4 outputs vs time
- Fig. 6 Human stress response ( $H_S$ ) vs assembly complexity ( $C$ ): (a) non-linear regression model, and (b) residual plots
- Fig. 7 Human stress response ( $H_S$ ) vs total defects ( $D_{tot}$ ) for single variant production: (a) linear regression model, and (b) residual plots
- Fig. 8 Average human stress response ( $\overline{H_S}$ ) vs average total defects ( $\overline{D_{tot}}$ ) for small batches of product variant: (a) linear regression model, and (b) residual plots
- Fig. 9 HRC-QWAT for (a) single variant production and (b) small-batch variant production

919

920  
921

### Table Caption List

Table 1	Components of the six electronic board variants (V1-V6)
Table 2	Participants' characteristics
Table 3	Difficulty of component handling attributes. Adapted from [63]
Table 4	Difficulty of component connection attributes. Adapted from [63]
Table 5	Handling complexity ( $h_p$ ) of components and connection complexity ( $c_{pr}$ ) of components with the breadboard in the six electronic board variants (V1-V6)
Table 6	Complexities of the six electronic boards (V1-V6)
Table 7	Experimental values of total defects ( $D_{tot}$ ) and human stress response ( $H_S$ ) recorded in each trial
Table 8	Number of defects classified into in-process (D1) and offline (D2) defects for the six assembled products
Table 9	Poisson regression output for total defects ( $D_{tot}$ ) vs assembly complexity ( $C$ ). Model is in the form $D_{tot} = (k_1 \cdot C)^2$
Table 10	Non-linear regression output for human stress response ( $H_S$ ) vs assembly complexity ( $C$ ). Model is in the form $H_S = k_2 \cdot C^{k_3}$
Table 11	Linear regression output for human stress response vs total defects for single variant production and small-batch variant production
Table 12	Examples of critical situations detected by the HRC-QWAT

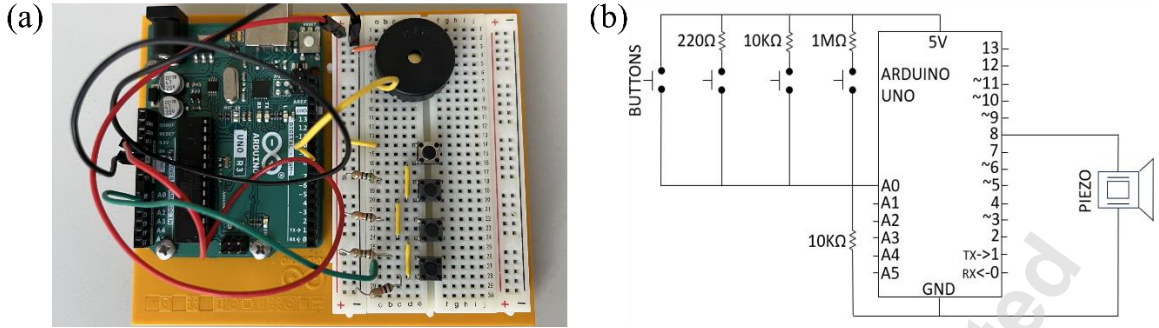
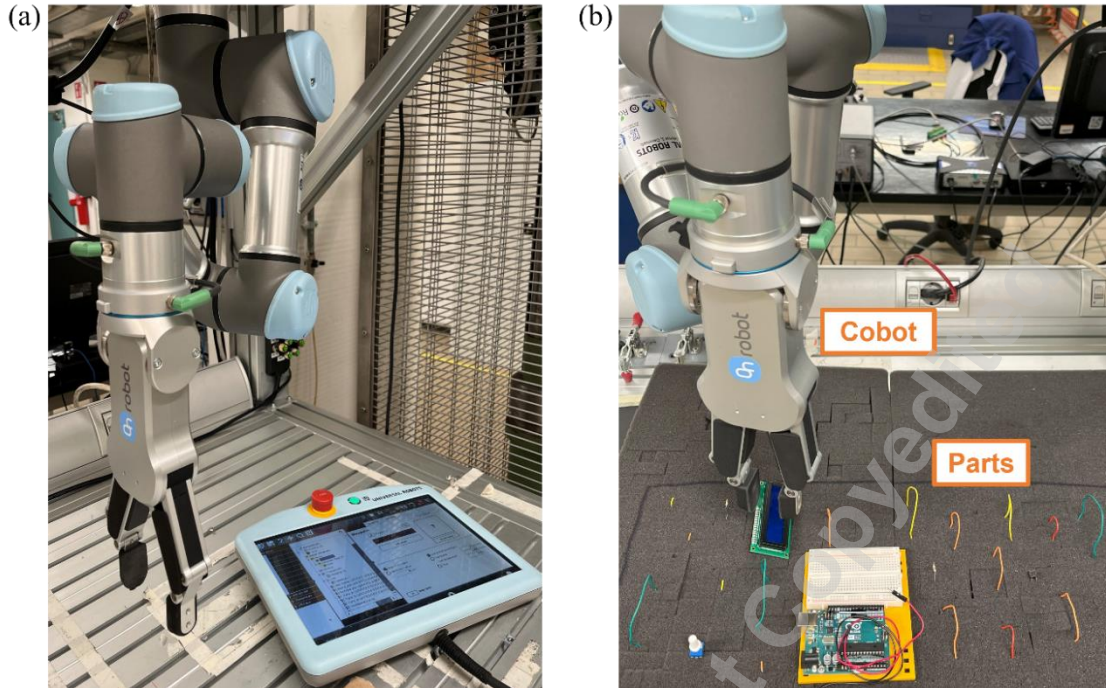


Fig. 1 Example of an assembled electronic board (variant V3): (a) final product assembled and (b) circuit diagram

922  
923  
924  
925

Accepted Manuscript Not Copied



926  
927  
928

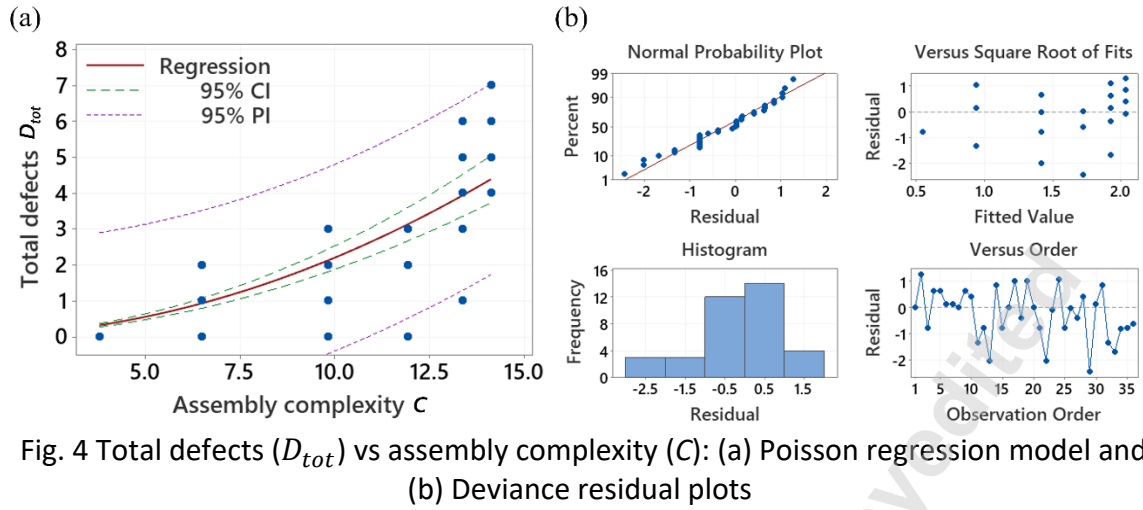
Fig. 2 Collaborative assembly workstation showing (a) the single-armed UR3e cobot equipped with the OnRobot RG6 gripper and (b) product components

Accepted Manuscript

$$AM = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Fig. 3 **AM** matrix of variant V3

929  
 930  
 931



932  
933  
934  
935

Accepted Manuscript Not Copyable

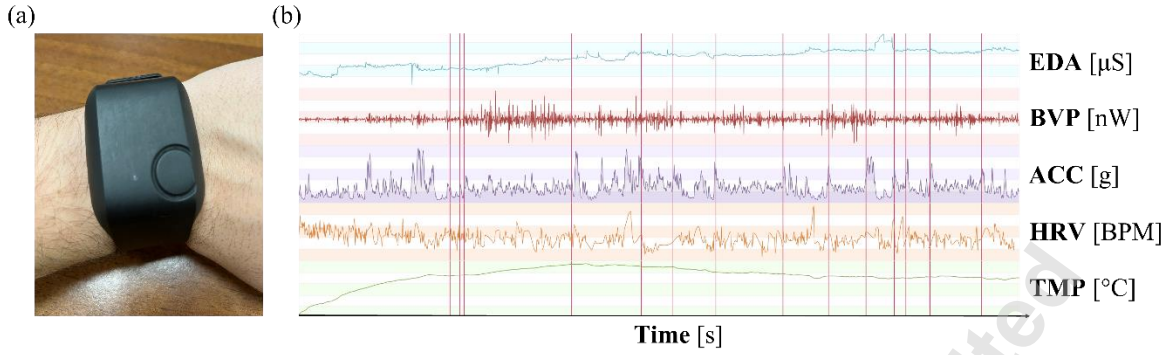


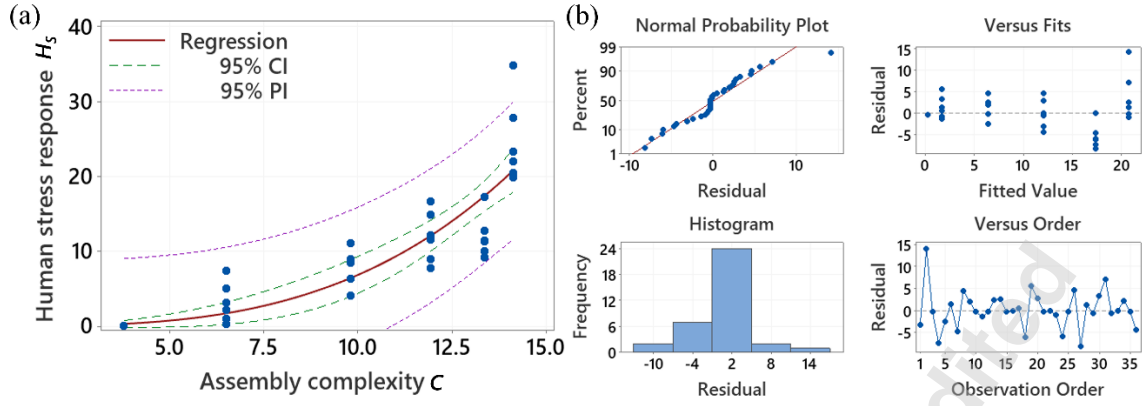
Fig. 5 (a) Empatica E4 wristband (b) Empatica E4 outputs vs time

936

937

938

Accepted Manuscript Not Copyable



939

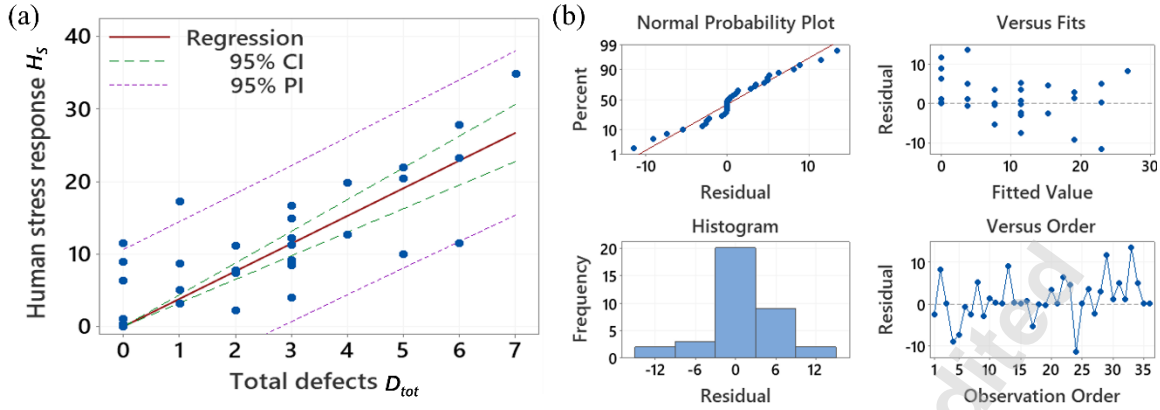
940

941

942

Fig. 6 Human stress response ( $H_s$ ) vs assembly complexity ( $C$ ): (a) non-linear regression model, and (b) residual plots

Accepted Manuscript Not Copyrigh



943

944

945

946

Fig. 7 Human stress response ( $H_S$ ) vs total defects ( $D_{tot}$ ) for single variant production:  
(a) linear regression model, and (b) residual plots

Accepted Manuscript Not Copyable

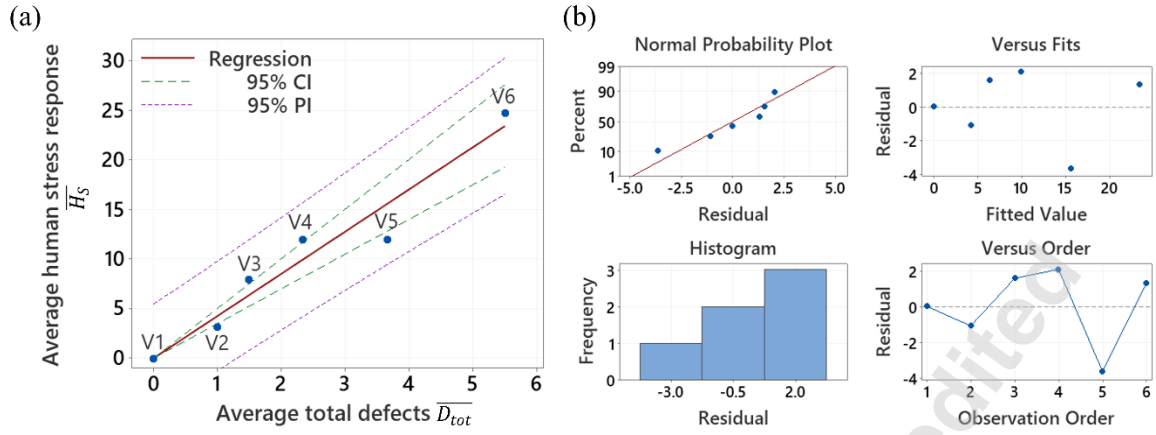
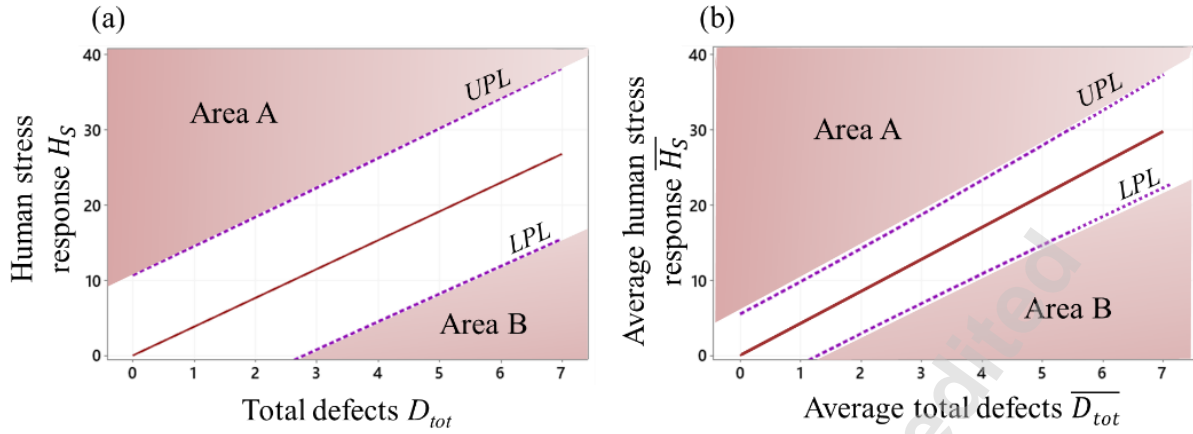


Fig. 8 Average human stress response ( $\overline{H}_S$ ) vs average total defects ( $\overline{D}_{tot}$ ) for small batches of product variant: (a) linear regression model, and (b) residual plots

947  
 948  
 949  
 950

Accepted Manuscript Not Copied



951  
952  
953

Fig. 9 HRC-QWAT for (a) single variant production and (b) small-batch variant production

Accepted Manuscript Not Certified

954

Table 1 Components of the six electronic board variants (V1-V6)

	V1	V2	V3	V4	V5	V6
Breadboard (BB)	1	1	1	1	1	1
Long wires (LW)	-	1	2	8	9	13
Short wires (SW)	1	3	5	3	6	4
Resistors (R)	1	1	4	6	2	2
Pushbuttons (PB)	-	2	4	-	2	1
LED (L)	1	1	-	1	-	-
Phototransistor (F)	-	-	-	3	-	-
Potentiometer (PT)	-	-	-	-	1	1
Piezo (PZ)	-	-	1	-	-	-
LCD (LCD)	-	-	-	-	-	1
Battery snap (BS)	-	-	-	-	1	-
DC Motor (M)	-	-	-	-	1	-
H-bridge (HB)	-	-	-	-	1	-
N° of components	4	9	17	22	24	23

955

956

957

Table 2 Participants' characteristics

Participant	Age	Gender	Domain knowledge of HRC	Domain knowledge of electronic board assembly
P1	21	Female	Intermediate	Intermediate
P2	21	Male	Intermediate	Expert
P3	22	Male	Expert	Expert
P4	21	Male	Intermediate	Expert
P5	27	Male	Intermediate	Expert
P6	23	Male	Intermediate	Intermediate

958

959

Accepted Manuscript Not Copyedited

960

Table 3 Difficulty of component handling attributes. Adapted from [63]

Attribute $i$	Description	Handling difficulty $d_h$
A - Size and weight (One of the following)	Very small - requires handling aids	1.5
	Easy - requires one hand only	1
	Large and/or heavy - requires more than one hand or aid	1.5
	Large and/or heavy- requires hoist or more than one person	2
B - Handling difficulty (All that apply)	Delicate	0.4
	Flexible	0.6
	Sticky	0.5
	Tangible	0.8
	Severely nest	0.7
	Sharp/abrasive	0.3
	Untouchable	0.5
	Gripping problem/slippery	0.2
C - Alpha Symmetry (One of the following)	Automatic handling - no difficulty	0
	Symmetrical - no orientation required	0
	Easy to orient - end to end	0.1
D - Beta Symmetry (One of the following)	Difficult to orient - end to end	0.5
	Rotational orientation is not required	0
	Easy to orient - end to end	0.2
	Difficult to orient - end to end	0.4

961

962

963

964

Table 4 Difficulty of component connection attributes. Adapted from [63]

Attribute $j$	Description	Connection difficulty $d_c$
E - Component placing (One of the following)	Self-holding	1
	Holding down required	2
	Self-securing	1.3
	Screwing	4
F – Component fastening (One of the following)	Riveting	4
	Bending	4
	Mechanical deformation	4
	Soldering or welding	6
	Adhesive	5
G – Direction (One of the following)	Straight line from above	0
	Straight line not from above	0.1
	Not straight line and/or bending is required	1.6
H – Insertion (One of the following)	Single	0
	Multiple	0.7
	Simultaneous multiple insertions	1.2
I – Restricted vision (One of the following)	Visible	0
	Not visible	1
J – Difficult to align (One of the following)	No	0
	Yes	0.7
K – Resistance to insertion (One of the following)	No	0
	Yes	0.6

965

966

967

968 Table 5 Handling complexity ( $h_p$ ) of components and connection complexity ( $c_{pr}$ ) of  
 969 components with the breadboard in the six electronic board variants (V1-V6)

Component	$h_p$	$c_{pr}$
Breadboard (BB)	1.7	-
Long wires (LW)	1.8	3.7, 5.3, 6.3
Short wires (SW)	2.3	3.7, 5.3
Resistors (R)	1.8	3.8
Pushbuttons (PB)	1.9	4.2
LED (L)	1.9	4.2
Phototransistor (F)	1.9	4.2
Potentiometer (PT)	1.7	5.8
Piezo (PZ)	1.7	3.7
LCD (LCD)	3.0	6.4
Battery snap (BS)	1.8	3.7
DC Motor (M)	1.8	3.7
H-bridge (HB)	1.9	4.2

970

971

972

Accepted Manuscript Not Copyedited

973

Table 6 Complexities of the six electronic boards (V1-V6)

	V1	V2	V3	V4	V5	V6
$C_1$	1.64	3.12	5.35	6.59	7.49	6.97
$C_2$	2.90	5.89	10.03	13.39	15.83	18.24
$C_3$	0.75	0.57	0.45	0.40	0.37	0.39
$C$	3.80	6.50	9.83	11.95	13.37	14.12

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

Accepted Manuscript Not Copyedited

995 Table 7 Experimental values of total defects ( $D_{tot}$ ) and human stress response ( $H_S$ )  
 996 recorded in each trial

Participant	Variant	$C$	$D_{tot}$	$H_S$
1	V4	11.95	3	8.97
1	V6	14.12	7	34.87
1	V1	3.8	0	0.00
1	V5	13.37	5	10.00
1	V3	9.83	3	4.02
1	V2	6.5	1	3.13
2	V5	13.37	4	12.70
2	V4	11.95	3	16.65
2	V3	9.83	3	8.46
2	V6	14.12	5	20.45
2	V2	6.5	0	0.33
2	V1	3.8	0	0.00
3	V3	9.83	0	8.95
3	V6	14.12	6	23.27
3	V1	3.8	0	0.00
3	V4	11.95	3	12.16
3	V2	6.5	2	2.23
3	V5	13.37	3	11.30
4	V2	6.5	2	7.35
4	V4	11.95	3	14.90
4	V1	3.8	0	0.00
4	V3	9.83	0	6.35
4	V6	14.12	4	19.84
4	V5	13.37	6	11.45
5	V1	3.8	0	0.00
5	V3	9.83	2	11.12
5	V5	13.37	3	9.21
5	V6	14.12	5	22.01
5	V4	11.95	0	11.55
5	V2	6.5	1	5.00
6	V6	14.12	6	27.88
6	V2	6.5	0	1.04
6	V5	13.37	1	17.31
6	V3	9.83	1	8.75
6	V1	3.8	0	0.00
6	V4	11.95	2	7.75

997

998 Table 8 Number of defects classified into in-process (D1) and offline (D2) defects for the  
 999 six assembled products

Variant	Incorrect Component		Misplaced Component		Unpicked Component	Slipped Component	Defective Component		Improperly Inserted Component	
	D1	D2	D1	D2	D1	D1	D1	D2	D1	D2
V1	0	0	0	0	0	0	0	0	0	0
V2	0	0	1	1	3	0	0	0	0	1
V3	0	0	5	2	3	0	0	0	0	1
V4	0	0	4	3	4	0	0	0	3	0
V5	0	0	6	3	11	2	0	0	0	0
V6	0	0	11	11	10	0	0	0	1	0
Total	0	0	27	20	31	2	0	0	4	2

1000  
 1001

Accepted Manuscript Not Certified

1002 Table 9 Poisson regression output for total defects ( $D_{tot}$ ) vs assembly complexity ( $C$ ).  
 1003 Model is in the form  $D_{tot} = (k_1 \cdot C)^2$

$k_1$	$SE(k_1)$	Coefficient $p$ -value	Deviance $R^2$	Goodness-of-Fit Tests	
0.144	0.008	<0.0005	99.29%	Deviance Test $p$ -value	0.557
				Pearson Test $p$ -value	0.933

1004

1005

Accepted Manuscript Not Copyedited

1006 Table 10 Non-linear regression output for human stress response ( $H_S$ ) vs assembly  
1007 complexity ( $C$ ). Model is in the form  $H_S = k_2 \cdot C^{k_3}$

$k_2$	$SE(k_2)$	95% CI for $k_2$	$k_3$	$SE(k_3)$	95% CI for $k_3$	$S$
0.004	0.006	$(4 \cdot 10^{-5}, 0.076)$	3.222	0.594	(2.086, 5.025)	4.284

1008  
1009

Accepted Manuscript Not Copyedited

1010 Table 11 Linear regression output for human stress response vs total defects for single  
 1011 variant production and small-batch variant production

	Model	$k_4$	$SE(k_4)$	Coefficient $p$ -value	$R^2$	$R^2$ pred.	$S$
Single variant production	$H_S = k_4 \cdot D_{tot}$	3.821	0.278	<0.0005	84.38%	82.99%	5.243
Small-batch variant production	$\bar{H}_S = k_4 \cdot \bar{D}_{tot}$	4.257	0.294	<0.0005	97.67%	95.64%	2.127

1012

Accepted Manuscript Not Copyedited

1013

Table 12 Examples of critical situations detected by the HRC-QWAT

HRC-QWAT	Observed values	Possible root cause
Single variant production	$(D_{tot}, H_S)=(2,35)$ Area A (cf. Fig. 9(a))	Abnormal stress experienced by the operator
Small-batch variant production	$(\overline{D}_{tot}, \overline{H}_S)=(6,10)$ Area B (cf. Fig. 9(b))	Wrong/faulty components undetected by the operator during the production process

1014

1015

Accepted Manuscript Not Copyedited