A structured methodology to support human–robot collaboration configuration choice

Original
A structured methodology to support human–robot collaboration configuration choice / Gervasi, R.; Mastrogiacomo, L.; Maisano, D.; Antonelli, D.; Franceschini, F.. - In: PRODUCTION ENGINEERING. - ISSN 0944-6524. - STAMPA. - 16:4(2022), pp. 435-451. [10.1007/s11740-021-01088-6]

Availability:
This version is available at: 11583/2957021 since: 2022-07-15T13:44:23Z

Publisher:
Springer Science and Business Media Deutschland GmbH

Published
DOI:10.1007/s11740-021-01088-6

(Article begins on next page)

A structured methodology to support Human-Robot Collaboration configuration choice

Gervasi, Riccardo[1)], Mastrogiacomo, Luca[1)], Maisano, Domenico Augusto[1)], Antonelli, Dario[1)] and Franceschini, Fiorenzo[1)]

[1)] Department of Management and Production Engineering (DIGEP), Politecnico di Torino, Turin, Italy

ABSTRACT

Human-Robot Collaboration (HRC) is a gradually consolidating paradigm of the modern industry which combines human and robot skills to make production more flexible. Since the effective implementation of HRC requires a careful analysis of different aspects, related both to robots and humans, there is a real need for a structured methodology to support it.

A previous work proposed a multi-dimensional framework to analyze several HRC aspects of a collaborative task. However, identifying the configuration that better exploits the HRC potential is not always trivial, especially among multiple alternative solutions. In addition, the priority levels (weights) assigned to the individual sub-dimensions of the framework, which identify specific design strategies, do not appear explicitly. The goal of this paper is to address these gaps, expanding the previous methodology and proposing the introduction of a Multiple-Criteria Decision Analysis (MCDA) method (i.e., ELECTRE-II). The inclusion of a MCDA method allows designers to: (i) express importance weights for each sub-dimension of the framework, and (ii) generate a preference ranking through a structured comparison of alternative HRC configurations. The description is supported by a real industrial application in the automotive field, in which four alternative HRC configurations are analyzed by a team of experts providing a holistic analysis.

Keywords: Human-Robot Collaboration, Multi-dimensional reference framework, Industry 4.0, Automotive industry, Assembly.

1. INTRODUCTION

Sharing of workspace and physical interaction between humans and robots in manufacturing processes are no longer a futuristic utopia, but a reality that has been consolidating in recent years. Unlike traditional robotic systems, collaborative robots represent a promising solution to meet the increasingly pressing demand for production based on the so-called “mass customization” [1, 2].

Being flexible, easily reconfigurable, efficient and adaptive, collaborative robots (or cobots) are one of the enabling technologies of Industry 4.0 [1, 3]. At the same time, they provide an important opportunity for technological development in many areas where robotics is almost unfamiliar [4, 5].

The main idea of Human-Robot Collaboration (HRC) is combining the abilities of humans with those of robots. On the one hand, humans have innate flexibility, intelligence, dexterity, and problem-solving skills, on the other hand, robots provide precision, power, and repeatability [6]. The implementation of HRC introduces several issues related mainly to safety [7–10], robot programming [4, 11], task organization [12, 13], and human-related aspects (e.g., physical and cognitive workload, ergonomics, usability, and acceptance) [14].

For an effective implementation of HRC, it is necessary to consider the whole of the above mentioned aspects [15, 16]. Many of the state-of-the-art HRC methodologies mainly focus on (i) a portion of these aspects [8, 17, 18] and/or (ii) specific tasks or situations [9, 13, 19, 20]. The attempt to build a general evaluation framework for HRC, able to consider all its aspects, seems to be less explored.

In a previous work, Gervasi, *et al.* [21] proposed a conceptual reference framework to evaluate HRC tasks, based on a plurality of dimensions. This framework helps to provide a holistic view of the HRC problem, allowing for a complete description and representation of a collaborative task. The framework also allows to compare different HRC solutions, highlighting the aspects on which they differ [22]. However, identifying the solution that better exploits the HRC potential is often not so straightforward, especially when multiple alternatives are present. Moreover, the importance weight of the individual sub-dimensions of the framework, which identify specific design strategies, do not appear explicitly.

In order to address this gap, this paper introduces a novel methodology to guide a team of experts in choosing between different design solutions for a collaborative task, allowing to maintain a holistic view of the problem. The main goals of this paper can be summarized as follows: (i) emphasizing the need for a complete view in collaborative task design and analysis for successful implementation of HRC; (ii) showing the use of the framework to evaluate and compare different configurations of an industrial task; (iii) guiding designers’ to express specific design strategies through importance weight assignment to sub-dimensions of the framework; and (iv) proposing a structured methodology to support a team of experts in comparing and choosing between different alternatives of HRC configurations of an industrial task.

The paper is organized as follows. Section 2 presents an overview of the literature on the HRC problem. Section 3 recalls the HRC evaluation framework proposed by Gervasi, *et al.* [21], highlighting its features. Section 4 describes and deeply analyzes four HRC configurations of an industrial HRC application in the automotive sector. In Section 5, a Multiple-Criteria Decision Analysis (MCDA) method (i.e., ELECTRE-II) is applied to support importance weight assignment to the framework sub-dimensions and generate a ranking of the four HRC solutions, identifying the most preferable one. Section 6 discusses the results obtained and the implications of the proposed methology. Finally, Section 7 contains conclusions and future research directions. Further details on the application of the ELECTRE-II method are contained in Appendix A and B sections.

2. LITERATURE REVIEW

HRC shares several aspects with Human-Robot Interaction (HRI), which is a more general research field dedicated to understanding, designing, and evaluating robotic systems to be used by or with humans [16]. HRC is focused on the implementation of collaborative robots, i.e., robots specifically designed to share space and also physically interact with humans.

The term "cobot" was first introduced by Colgate *et al.* [23] in 1996, presenting a robotic system able to manipulate objects in conjunction with a human operator. In this work, the collaboration was interpreted as a form of assistance, by constraining and guiding human movements in specific operations.

The ability to physically interact with a robot in an industrial setting subverts the classic paradigm of separation of workspaces between humans and industrial robots. However, although the removal of this limitation opens new horizons towards the management of production processes, it introduces new safety hazards for the operator [8]. The growing attention to these issues and the development of robotic assistants in the industrial environment has led to an evolution of safety standards.

ISO 10218-1 and ISO 10218-2 provide guidelines on the implementation of industrial robots and workspace design, identifying a list of safety hazards. The subsequent ISO/TS 15066 has expanded the possibilities for HRC, proposing different collaborative modes and allowing the implementation of higher levels of robot autonomy in proximity to humans.

In addition to the safety aspects of HRC, the literature over time has focused on task organization and robot programming methods. Task organization plays a central role due to the high influence on other aspects of HRC, such as performance, ergonomics, and workload [1, 24]. The classic robot programming approach is manual, where the user manually implements the actions to be performed by the robot using text-based or graphical languages [11]. In order to increase the intuitiveness of robot programming, new methods have been developed to instruct collaborative robots, such as Programming by Demonstration (PbD). PbD consists of instructing the cobot by showing it the sequence of operations it will have to reproduce. This technique thus allows the state of the cobot to be recorded while the operator guides it, physically or via a controller, along the operations to be performed [11].

In the HRC paradigm, human-related aspects have a key role in successful implementation. The introduction of collaborative robots in a manufacturing context has an impact on the human operators involved [25]. Affective state and cognitive processes can greatly influence the success of HRC and, consequently, task performance [26, 27]. Minimizing stresses from the work environment and interaction with the robot is therefore necessary to make HRC more effective [28, 29]. Another key aspect that requires special attention in the implementation of HRC is the acceptance of the technology by the operators [30, 31]. For the successful implementation of collaborative robots, it is important that the operators are involved in the implementation process, there is adequate communication of the change, and the workforce receives appropriate training [25].

The aspects that characterize HRC are multiple and diverse, requiring the fundamental contribution of several disciplines to understand, design, and evaluate HRC with a complete view. Some frameworks to support the implementation of HRC have been proposed recently.

Mateus *et al.* [1] presented a methodology to guide in the design of collaborative assembly tasks. Considering the robot's capabilities and aspects of physical ergonomics, safety, and execution time, it provides assembly sequences including work allocation.

Rajendran *et al.* [32] proposed a framework to support HRI user studies based on Robot Operating System (ROS) middleware. The main goal of the framework is to simplify the implementation and reproducibility of HRI user studies, while also providing metrics for real-time evaluations. At present, the framework focuses primarily on evaluating team performance and fluency, however the modularity of the framework allows for other metrics to be implemented.

Kopp et al. [33] conducted a study aimed at identifying success factors for introducing HRI in an industrial setting. The factors identified considered both technical and socio-cultural aspects. These factors were classified into a framework composed of two dimensions: the component of an HRI solution and the HRI introduction phase. The former consisted of four levels (human operator, cobot, work system, and enterprise and context), while the latter consisted of three levels (decision phase, implementation phase, and operation phase).

In the next section, the holistic evaluation framework for HRC proposed by Gervasi et al. [21] will be presented, highlighting its main features. and used to evaluate and compare different configurations of a collaborative task.

3. HRC EVALUATION FRAMEWORK

Gervasi, *et al*. [21] proposed a conceptual reference framework to evaluate HRC applications, considering a plurality of dimensions and their evaluation. The main features of this framework can be summarized as follows: (i) it provides a holistic view of the HRC problem, considering aspects related to the robot, the human operator, the working system, and the working context; (ii) it is applicable for the representation and evaluation of different types of collaborative tasks (e.g., assembly, polishing, grinding, welding, inspection, and co-manipulation) due to its generality; (iii) it allows the comparison between different solutions of the same collaborative task on the multiple aspects of HRC.

Table 1 summarizes the structure of the HRC framework with proposed evaluation methods for each (sub-)dimension. For additional details, refer to Gervasi, *et al.* [21]. A brief description of the (sub-)dimensions of the framework of interest follows below:

- *Autonomy* represents the robot capability to sense the surroundings, plan and act according to the environment and other entities. In the HRC context, high autonomy does not imply the exclusion of the human but enables a more advanced and deeper interaction, as in human-human interaction [16, 34]. Note that this view is opposite to that of automation, where autonomy is generally interpreted as the extent to which a system can perform a task without human intervention.
- *Information Exchange* represents the way information is exchanged between robot and human. Communication is the basis of any type of interaction between entities and is used to give commands, transmit information, and notify status. It is composed of two sub-dimensions: (i) *Communication format*, which refers to the senses involved in the communication; (ii) *Communication medium*, which refers to the way communication takes place.
- *Team Organization* concerns the organizational configuration of the agents involved in the collaboration. It includes two sub-dimensions: (i) *Structure of the team*, which refers to the number of robots and humans in the team; (ii) *Role of members*, which specifies the role of each team member.

- *Adaptivity and Training* concerns adaptivity and instruction of the robot as well as human training and it is characterized by three sub-dimensions: (i) *Robot adaptivity*, representing the ability to accomplish a given task despite unexpected situations; (ii) *Robot training method*, referring to the methods for instructing the robot to perform a certain task; (iii) *Operator training*, which indicates the effort in training the operators involved in a collaborative task.
- The *Task* dimension consists of four sub-dimensions: (i) *Field of application*, referring to the context in which the task takes place; (ii) *Task organization*, referring to the assignment of individual operations to each team member; (iii) *Performance*, referring to the evaluation of the collaborative outcome (e.g., effectiveness, efficiency, etc.), which may vary according to the collaborative task; (iv) *Safety*, concerning the identification of risks/hazards introduced by the collaborative task and the related safety measures implemented.
- The *Human Factors* dimension analyzes the interactions among human and robot to optimize human well-being and overall system performance [35]. It is composed of five sub-dimensions: (i) *Workload*, referring to both the physical and mental effort of human operators during a task; (ii) *Trust*, representing the propensity of an agent to achieve a certain goal in a situation characterized by uncertainty and vulnerability induced by the task, the cobot, and its reliability [25]; (iii) *Robot morphology*, evaluating the morphological and design features of the collaborative robot; (iv) *Physical ergonomics*, concerning the anatomical, anthropometric and biomechanical characteristics of humans in relation to physical activity and the evaluation of related hazards; (v) *Usability*, referring to the extent to which human-robot collaboration is effective, efficient, and satisfying to the user in order to achieve certain objectives.
- *Ethics* represents the common understanding of the principles that constrain and guide human behavior [36]. This dimension includes two sub-dimensions: (i) *Social impact*, concerning the consequences of introducing an HRC system within a community; (ii) *Social acceptance*, referring to the perception of the system within a community.
- *Cybersecurity* represents the process of protecting sensitive data by preventing, detecting, and responding to cyberattacks [37]. It is composed of five sub-dimensions: (i) *Identification*, concerning the actions related to understanding policies, cybersecurity risks, and priorities that can be relevant to manage cybersecurity risks; (ii) *Protection*, activities related to the development and implementation of safeguards to protect infrastructure services; (iii) *Detection*, involving training and other activities related to the detection of cyberattacks; (iv) *Response*, involving (re)action following the detection of a certain cyberattack; (v) *Recovery*, representing activities to recover from a certain cyberattack.

Table 1 – Summary of the HRC conceptual framework with its dimensions, sub-dimensions, and evaluation methods [21].

Dimension	Sub-dimension	Evaluation method	Scale levels
Autonomy	-	LORA [17]	(L0) Manual – (L1) Teleoperation – (L2) Assisted Teleoperation – (L3) Batch Processing – (L4) Decision Support – (L5) Shared Control with Human Initiative – (L6) Shared Control with Robot Initiative – (L7) Executive Control – (L8) Supervisory Control – (L9) Full Autonomy
Information Exchange	Communication medium	4-level scale	(L0) No senses involved – (L1) A sense between sight, hearing, and touch is involved – (L2) Two senses between sight, hearing, and touch are involved – (L3) Sight, hearing, and touch are involved
	Communication format	4-level scale	(L0) No means – (L1) Only control panel/displays – (L2) A human-natural communication mean is implemented – (L3) At least two human-natural communication means are implemented
Team Organization	Team structure	Categorical scale	List of robots and humans involved.
	Member role	3-level scale	(L0) Executor – (L1) Assistant – (L2) Master
Adaptivity and Training	Robot adaptivity	4-level scale [38]	(L0) No adaptivity – (L1) No flexible adaptivity – (L2) Adaptivity – (L3) Adaptivity with respect to human
	Robot training method	3-level scale	(L0) Only manual programming – (L1) Automatic programming are implemented – (L2) Automatic programming methods based on natural communication are implemented
	Operator training	4-level scale	(L0) Very Heavy – (L1) Heavy – (L2) Medium – (L3) Light
Task	Field of application	Categorical scale	Description of the application context.
	Task organitation	List of operations	-
	Performance	4-level scale	(L0) Low – (L1) Medium – (L2) High – (L3) Very High
	Safety	Risk Assessment [39, 40]	(L0) Low – (L1) Medium – (L2) High – (L3) Very High

Table 1 – (continued)

Dimension	Sub-dimension	Evaluation method	Scale levels
Human Factors	Workload	NASA-TLX [41]	(L0) Very High – (L1) High – (L2) Medium – (L3) Low
	Trust	Trust Scale questionnaire [25]	(L0) Low – (L1) Medium – (L2) High – (L3) Very High
	Robot morphology	Categorical scale [42]	Anthropomorphic – Zoomorphic – Functional
	Physical ergonomics	EAWS [43]	(L0) Red – (L1) Yellow – (L2) Green
	Usability	SUS [44, 45]	(L0) Not acceptable – (L1) Marginal – (L2) Acceptable
Ethics	Social impact	3-level scale	(L0) Heavy – (L1) Medium – (L2) Light
	Social acceptance	Brohl TAM [18]	(L0) Low – (L1) Medium – (L2) High – (L3) Very High
Cybersecurity	Identification	Dedeke framework [46]	(L0) Partial – (L1) Risk informed – (L2) Repeatable – (L3) Adaptive
	Protection		(L0) Partial – (L1) Risk informed – (L2) Repeatable – (L3) Adaptive
	Detection		(L0) Partial – (L1) Risk informed – (L2) Repeatable – (L3) Adaptive
	Response		(L0) Partial – (L1) Risk informed – (L2) Repeatable – (L3) Adaptive
	Recovery		(L0) Partial – (L1) Risk informed – (L2) Repeatable – (L3) Adaptive

4. DATA COLLECTION AND METHODOLOGY

The proposed framework can be used to analyze alternative configurations related to a specific HRC application. Precisely, the application of interest takes place in an automotive company and concerns the specific task of assembling a gearbox component, called "parking pawl" [22]. A team of experts, interacting with managers and operators, created four HRC alternative configurations, which are described in the following sub-sections. These configurations were devised by introducing variations on several design aspects that could affect various dimensions of the HRC, while still maintaining a similar implementation cost. The main design constraint involved the task execution time, which had to be under 5 minutes.

4.1. First configuration (HRC1)

The workstation for the assembly task of interest is managed by three agents: a robotic system and two human operators (i.e., a logistics operator and a process operator) sharing the same workspace without any physical or virtual safety barrier.

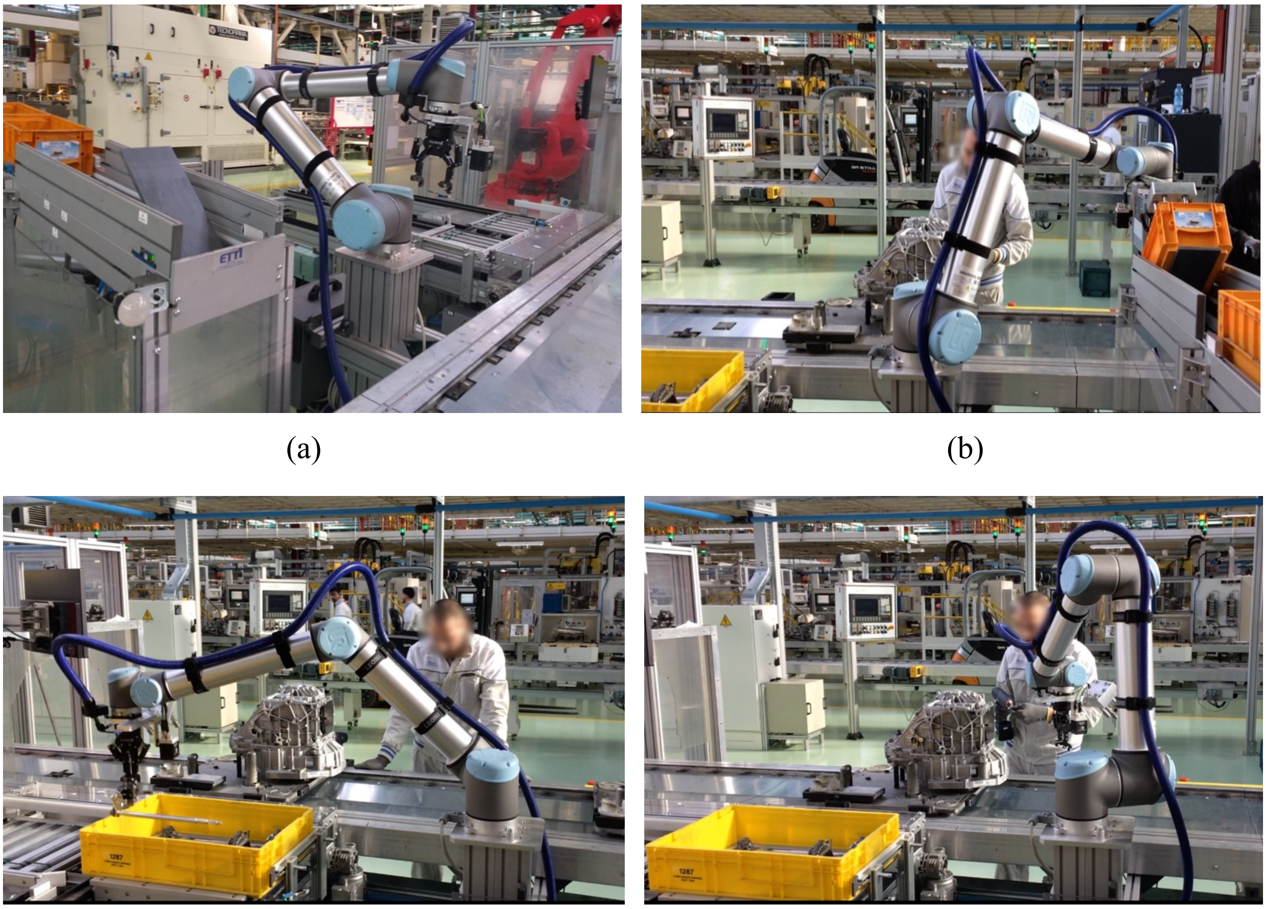
The robot system is a single-arm collaborative robot UR10/CB3 [47], equipped with three end-effectors installed on the robot flange, i.e., (i) an electromagnetic gripper to pick screws from a box, (ii) a vision system SensoPart Visor V20 2D, and (iii) a collaborative gripper Robotiq 2F-85.

The parking-pawl assembly task can be decomposed into four main operations:

1. The logistics operator sets up the workpieces in appropriate boxes, checking their correct position (Figure 1a).
2. The robot picks six screws from the workpiece box, through the electromagnetic gripper, and hands them to the operator (Figure 1b).
3. Using the gripper, the robot grabs the parking pawl and hands it to the operator in an ergonomic position (Figure 1c).
4. The operator inserts the parking pawl into the gearbox and tightens it using a screwdriver (Figure 1d).

Table 2 summarizes the evaluations based on the proposed HRC framework (in Section 2) for the four HRC configurations. Below, the results related to HRC1 are illustrated in detail [22].

Autonomy of the robot was rated L3 (“Batch Processing”) since the robot supports the operator during the task and is able to sense the environment, thanks to the vision system and force sensors, but task planning is exclusive to the human.



(a) (b) (c) (d)

Figure 1 – Main operations of the parking-pawl assembly task: (a) components setup, (b) screws feeding, (c) pawl feeding, and (d) pawl screwing.

Communication with the robot takes place through (i) push buttons, to command the robot to release workpieces, and (ii) a teach pendant, which displays the robot's status. For this reason, *Communication medium* and *Communication format* were rated L2 and L1, respectively.

Robot adaptivity was rated L1, since the robot is equipped with a vision system to identify the location of the parking pawl and it is able to adapt its position and movement to make an optimal gripping. *Robot training method* was rated L1, as the robot can be instructed using a teach pendant, which represent an automatic programming method. Since the training course mainly focuses on task operations, safety settings and the use of the teach pendant, *Operator training* was rated L2 (Medium).

Based on managers' feedback, *Performance* was rated L2 (High) mainly due to the cycle time being between 3 and 4 minutes and a process defectiveness within constrains. *Safety* was assessed through a risk-assessment based on the list of hazards contained in ISO 10218-2 standard. The risk-assessment was based on the risk matrix proposed in ISO/TR 14121-2, in which the severity and probability of occurrence of harm were considered (Table 3). The assessment considered the risk reduction due to

Table 2 – Evaluation summary of the four HRC configurations

Dimension	Sub-dimension	Configuration rating			
		HRC1	**HRC2**	**HRC3**	**HRC4**
Autonomy	-	L3 (Batch processing)	L3	L3	L3
Information Exchange	Communication medium	L2	L2	L2	L3
	Communication format	L1	L1	L2	L3
Team Organization	Team structure	2 Humans, 1 Robot	2 Humans, 1 Robot	2 Humans, 1 Robot	2 Humans, 1 Robot
	Member role	Human (1) L2 (Master) Human (2) L1 (Assistant) Robot L0 (Executor)	Human (1) L2 (Master) Human (2) L1 (Assistant) Robot L0 (Executor)	Human (1) L2 (Master) Human (2) L1 (Assistant) Robot L0 (Executor)	Human (1) L2 (Master) Human (2) L1 (Assistant) Robot L0 (Executor)
Adaptivity and Training	Robot adaptivity	L1	L1	L1	L3
	Robot training method	L1	L1	L1	L3
	Operator training	L2 (Medium)	L2	L2	L1 (High)
Task	Field of application	Manufacturing (automotive)	Manufacturing (automotive)	Manufacturing (automotive)	Manufacturing (automotive)
	Performance	L2 (High)	L3 (Very High)	L2	L3
	Safety	L3 (Very High)	L2 (High)	L3	L3
Human Factors	Workload	L2 (Medium)	L2	L1 (High)	L2
	Trust	L3 (Very High)	L2 (High)	L3	L3
	Robot morphology	Functional (Single arm)	Functional (Single arm)	Functional (Single arm)	Functional (Single arm)
	Physical ergonomics	L2 (Green)	L2	L2	L2
	Usability	L1 (Marginal)	L1	L1	L2 (Acceptable)
Ethics	Social impact	L1 (Medium)	L1	L1	L1
	Social acceptance	L2 (High)	L2	L2	L1 (Medium)
Cybersecurity	Identification	L1 (Risk informed)	L1	L1	L2 (Repeatable)
	Protection	L1 (Risk informed)	L1	L1	L2 (Repeatable)
	Detection	L1 (Risk informed)	L1	L1	L2 (Repeatable)
	Response	L1 (Risk informed)	L1	L1	L2 (Repeatable)
	Recovery	L1 (Risk informed)	L1	L1	L2 (Repeatable)

Table 3 – Risk matrix proposed in ISO/TR 14121-2.

	Severity of harm			
Probability of occurrence	(L3) Catastrophic	(L2) Serious	(L1) Moderate	(L0) Minor
(L3) Very likely	High (3)	High (3)	High (3)	Medium (2)
(L2) Likely	High (3)	High (3)	Medium (2)	Low (1)
(L1) Unlikely	Medium (2)	Medium (2)	Low (1)	Negligible (0)
(L0) Remote	Low (1)	Low (1)	Negligible (0)	Negligible (0)

the implementation of protective measures, such as safety functions of the cobot, influencing the probability of occurrence and severity of harms. Regarding mechanical hazards, the most likely risks are "impact", "cutting/severing" and "friction/abrasion", due to the possibility of contact with moving workpieces and the cobot. However, the severity of harm of each of these risks was rated “Moderate” (L1), as the cobot safety functions significantly reduces the damage and the possible contact regions were not vital organs. The other mechanical hazards (“drawing-in/trapping”, “entanglement”, “crushing”, “shearing”, and “stabbing/puncture”) and hazards of other categories were evaluated with a “Serious” (L2) severity but “Remote” (L0) or “Unlikely” (L1) probability of occurrence. Some hazards were assessed as “Not Available” (N/A) since potential harm was completely excluded. The final risk score obtained was 22/90 (less than 25% of the maximum score), leading to a "Very High" (L3) evaluation for *Safety* according to the scale suggested in the HRC framework [21].

Workload, *Trust*, and *Usability* were rated respectively L2 (Medium), L3 (Very High) and L1 (Marginal) based on managers’ and operators’ feedback. *Physical ergonomics* was rated L2 (Green), using the EAWS tool [43]. The task requires the handling of low load objects and the application of low forces while maintaining a non-fatiguing posture, hence involving a low biomechanical load on the operator. This is confirmed by the EAWS score of 15.5 (< 25), which indicates a low risk of biomechanical overload.

Social acceptance was rated L2 (High) based on managers’ and operators’ feedback, while *Social impact* was rated L1 (Medium), since the introduction of the collaborative robot implied a redeployment of personnel. Originally, the parking pawl assembly was performed at a dedicated off-line station. This operation was performed manually by one operator, on average for two shifts per day. With the introduction of the cobot, this task was integrated directly into the production line, resulting also in a redeployment of personnel.

Identification, Protection, Detection, Response, and *Recovery* were all rated L1 (Risk informed), given the presence of dedicated cybersecurity staff.

4.2. Second configuration (HRC2)

As in the first HRC configuration, the workstation is managed by three agents: a robotic system and two human operators. The robot is the same, but it is equipped with the following end-effectors: (i) an electromagnetic gripper, (ii) a vision system and (iii) a collaborative screwdriver. The major assembly operations are:

1. The logistics operator sets up the workpieces in appropriate boxes, checking their correct position.
2. Through the electromagnetic gripper, the robot picks six screws from the workpiece box and hands them to the operator.
3. The operator picks the parking pawl up, inserts it into the gearbox and places the screws in their respective seats.
4. The robot tightens the screws using the screwdriver end-effector.

The evaluations based on the proposed HRC framework for the HRC2 configuration are reported in Table 2 and recalled below.

Autonomy, *Information Exchange*, *Team Organization*, *Adaptivity and Training*, *Physical ergonomics*, *Workload*, *Usability, Ethics,* and *Cybersecurity* have not undergone any changes compared to the first HRC configuration.

Performance was rated L3 (Very High), since the use of the robot can significantly improve the precision and repeatability of the screwing operation. *Safety* was downgraded to L2 (High) compared to the first HRC configuration, due to the presence of the screwdriver end-effector, which increases the risks of “crushing” and “stabbing/puncture”. *Trust* was also downgraded to L2 (High), since the screwdriver end-effector may reduce the operator's perception of safety.

4.3. Third configuration (HRC3)

The task of this HRC configuration is identical to HRC1, except that the robot can receive instructions by the operator only through vocal commands. The operator is equipped with a microphone through which he can control the collaborative robot. The evaluations based on the proposed HRC framework for the HRC3 configuration are reported in Table 2 and recalled as follows.

Autonomy, *Communication medium*, *Team Organization*, *Adaptivity and Training*, *Task*, *Trust*, *Physical ergonomics*, *Usability*, *Ethics*, and *Cybersecurity* are not changed compared to the first HRC configuration.

Communication format was rated L2 (“A human-natural communication mean is implemented”), since vocal commands represent a natural and intuitive communication mean. However, *Workload* was downgraded to L1 (High), as the only use of vocal commands can be tiring for the operator in the long-term.

4.4. Fourth configuration (HRC4)

The task in the fourth HRC configuration is similar to HRC1, except that the robot can receive instructions by the operator not only through push buttons, but also through vocal commands and gestures. As in the HRC3 configuration, the operator is equipped with a microphone to give instructions to the robot. The operator can also give commands to the robot through a set of gestures. The gesture recognition is achieved through the Leap Motion Controller hand tracking sensor. In addition, the robot can adapt its movements and speed, taking into account the specific activity of the operator and his/her relative position.

The evaluations based on the proposed HRC framework for the HRC4 configuration are shown in Table 2 and recalled below.

Autonomy, *Team Organization*, *Safety*, *Workload*, *Trust*, *Physical ergonomics*, and *Social impact* have not undergone any changes compared to the first HRC configuration.

Communication medium and *Communication format* were both rated L3, since communication with the robot can be performed through push buttons, voice, and gestures, while information on the robot's status is displayed on the teach pendant.

Robot adaptivity was rated L3 since the robot is able to adapt its behavior to the specific activity of the operator and his/her relative position. Since this configuration also allows to program the robot though gestures, *Robot training method* was updated to L2. However, *Operator training* was downgraded to L1 (High), as the required training effort is significantly higher.

Performance was rated L3 (Very High) since the improved adaptability of the robot can help the operator to reduce errors during the task.

Usability was upgraded to L2 (Acceptable), as the operator can interact with the robot through different communication modalities, depending on the practical situation.

Social acceptance was downgraded to L1 (Medium), since some operators may be opposed to being constantly monitored in their position and movements, as required by the robot's adaptation system.

In order to protect the operator's sensible data, cybersecurity was significantly improved with respect to the previous configurations, introducing risk-informed policies and practices. For this reason, *Identification, Protection, Detection, Response*, and *Recovery* were all updated to L2 (Repeatable).

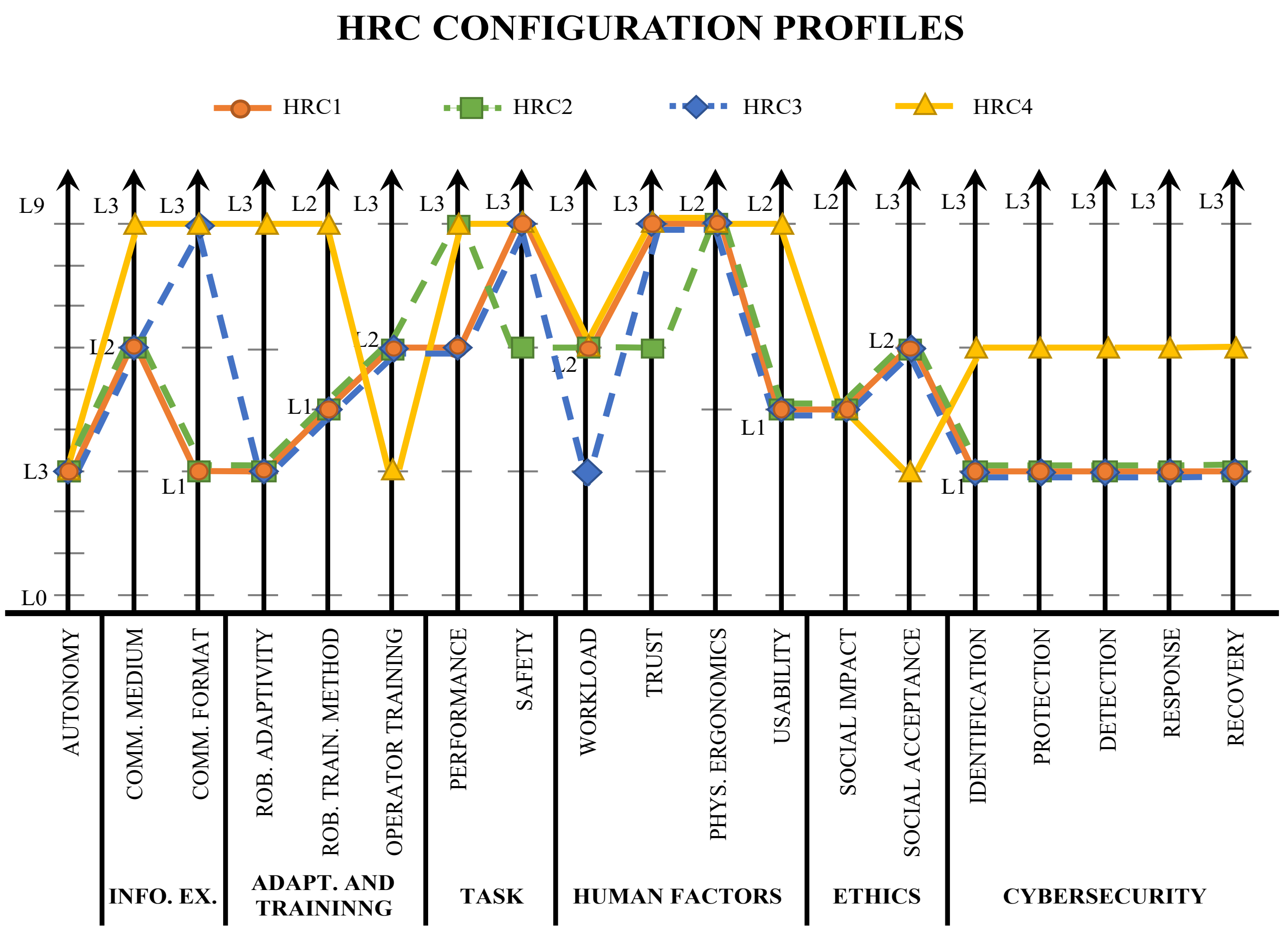


Figure 2 – Graphical comparison between the four HRC evaluation profiles: the first HRC configuration (orange), the second one (green), the third one (blue), and the fourth one (yellow).

5. RANKING OF HRC CONFIGURATIONS

The HRC evaluation framework allows to compare the four HRC configurations. As shown in Figure 2, the four configurations can be represented in the form of profiles, which represent the corresponding performance from the point of view of each (sub-)dimension. It can be noticed that all four evaluation profiles intersect each other, complicating the identification of the most preferable HRC configuration, i.e., the one that best meets the totality of the HRC sub-dimensions at the same time. It is therefore clear that the concept of *most preferable configuration* may depend on characteristic features of the HRC problem of interest, such as (i) the (sub-)dimensions considered and (ii) their different degree of importance, if they are not necessarily equally important.

A decision support method can be used to create a ranking of the HRC configurations. Although other methods are possible, ELECTRE II was chosen for the following reasons: (i) it allows to compare

Table 4 – Importance weights for each HRC dimension and sub-dimension.

Dimension	Dimension weight	Sub-dimension	Sub-dimension weight
Autonomy	9%	-	9.00%
Information Exchange	9%	Communication medium	2.97%
		Communication format	6.03%
Adaptivity and Training	14%	Robot adaptivity	7.00%
		Robot training method	2.38%
		Operator training	4.62%
Task	20%	Performance	6.60%
		Safety	13.40%
Human Factors	25%	Workload	8.75%
		Trust	3.75%
		Physical ergonomics	8.75%
		Usability	3.75%
Ethics	20%	Social impact	10.00%
		Social acceptance	10.00%
Cybersecurity	4%	Identification	0.80%
		Protection	0.80%
		Detection	0.80%
		Response	0.80%
		Recovery	0.80%

alternatives evaluated on multiple dimensions, (ii) it is suitable for dimensions rated on ordinal scales, and (iii) it is ideal for problems characterized by a number of heterogeneous dimensions [48].

Before using ELECTRE-II, a (cardinal) weight has to be assigned to each HRC (sub-)dimension. Among the various methods for assigning weights, it was decided to implement the Simos' procedure [49] mainly for the following reasons: (i) it is easy to use for decision makers, and (ii) it is often used in conjunction with ELECTRE methods [50]. Given the presence of a hierarchy between the dimensions of the HRC framework, a two-step procedure was implemented: (i) in the first step, a weight has been assigned to each main dimension using the Simos' procedure; (ii) subsequently, the weight of each dimension has been distributed among their respective sub-dimensions using again the Simos' procedure. For details on the application of the Simos' procedure to the problem of interest, see Appendix B. Table 4 contains the weights assigned to each HRC dimension and sub-dimension with the procedure described. It can be noted that the team of experts has given particular importance to *Human Factors*, *Task*, and *Ethics* since, in addition to safety and task performance, the social context and the well-being of operators can be essential to fully exploit the potential of collaboration.

Once each (sub-)dimension has been assigned a weight, the ELECTRE-II method can be applied. In the problem of interest, ELECTRE-II produced the following result: the most preferable HRC design solution turned out to be the fourth one, followed by the first one, the third one, and finally the second one (i.e., HRC4 ≻ HRC1 ≻ HRC3 ≻ HRC2, where the symbol "≻" denotes the strict preference relationship). For details on the application of the ELECTRE II method to the problem of interest, see Appendix B.

Since the assignment of weights to each sub-dimension has an inherently arbitrary component, it is possible to validate the results obtained by conducting a *sensitivity analysis*. The purpose of this analysis is to evaluate the robustness of the resulting collective ranking with respect to (relatively small) weight variations. For example, the possible variations in the collective ranking resulting from the ELECTRE-II application can be analyzed for three different weight combinations, as shown in Table 5 and Figure 3. The three weight combinations are obtained by artificially distorting the previous weights with a random noise ranging between -5% and 5%.

In spite of the variations between the weight combinations in use, the collective rankings resulting from the application of ELECTRE-II changed only slightly:

- weight combination No. 1 led to HRC4 ≻ HRC1 ≻ HRC3 ≻ HRC2;
- weight combination No. 2 led to HRC4 ≻ HRC3 ≻ HRC1 ≻ HRC2;
- weight combination No. 3 led to HRC4 ≻ HRC1 ≻ HRC3 ≻ HRC2.

Precisely, the collective rankings related to the first and third weight combination are identical to the one originally obtained, while the collective ranking related to the second weight combination differs only for a rank reversal between HRC1 and HRC3. Therefore, the solution provided by ELECTRE-II for the problem of interest appears robust.

6. DISCUSSION AND IMPLICATIONS

The HRC conceptual framework proposed by Gervasi, *et al.* [21] provides a multi-dimensional representation of a generic HRC task, taking into account the synergistic interaction of agents (robots and operators) and the specific application context. Additionally, this framework provides an environment for comparing and evaluating, from a collaborative point of view, various implementation solutions of a collaborative task. The holistic view proposed by the framework allows taking into account both technical and human aspects. However, one of the major difficulties of this

Table 5 – The three different weight combinations used for the sensitivity analysis of the result obtained through ELECTRE-II.

Sub-dimension	**Label**	**Weight combinations**		
		No. 1	**No. 2**	**No. 3**
Autonomy	e_1	9.52%	10.87%	12.88%
Communication medium	e_2	3.17%	2.99%	7.38%
Communication format	e_3	6.35%	5.05%	2.42%
Robot adaptivity	e_4	7.14%	4.85%	9.40%
Robot training method	e_5	2.38%	1.30%	1.03%
Operator training	e_6	4.76%	6.38%	0.78%
Performance	e_7	6.35%	3.16%	6.53%
Safety	e_8	12.70%	10.09%	13.07%
Workload	e_9	7.94%	4.66%	3.02%
Trust	e_{10}	3.97%	0.92%	2.03%
Physical ergonomics	e_{11}	7.94%	12.21%	4.05%
Usability	e_{12}	3.97%	0.94%	7.17%
Social impact	e_{13}	9.52%	8.04%	6.71%
Social acceptance	e_{14}	9.52%	7.09%	10.82%
Identification	e_{15}	0.95%	4.67%	4.06%
Protection	e_{16}	0.95%	5.61%	2.06%
Detection	e_{17}	0.95%	4.67%	0.51%
Response	e_{18}	0.95%	5.61%	4.06%
Recovery	e_{19}	0.95%	0.90%	2.01%

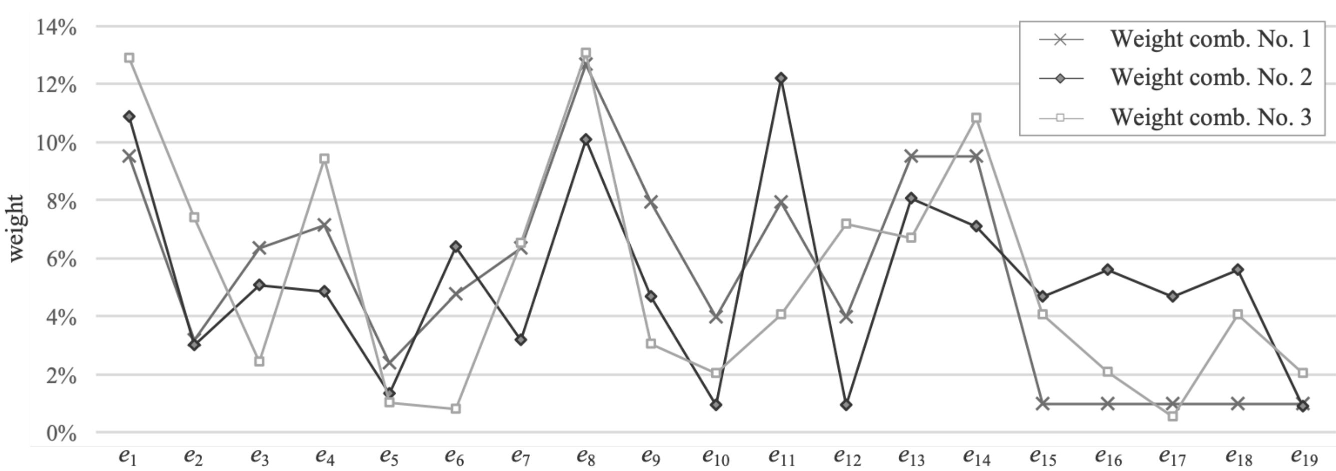


Figure 3 – Graphical representation of the three different weight combinations in Table 5.

tool is to identify which configuration(s) is/are the most preferable from a collaborative point of view. This is especially the case when there are several possible configurations, and a comparison of the resulting evaluation profiles reveals several "intersections" along the sub-dimensions of the framework. To overcome this limitation, this paper has introduced a methodology to guide a group of experts in choosing between different HRC configurations. In order to handle ordinal scale evaluations and the heterogeneous dimensions of the framework, the proposed methodology was based on a combination of the Simos' procedure and the ELECTRE-II method. Simos' procedure allows to assign importance weights to the framework dimensions through a simple indirect method. Through the sorting of cards containing the name of the framework dimensions, an importance ranking is generated from which importance weights are derived. The obtained distribution of the weights represents the design strategy implemented by the designers and can vary depending on the specific context of the problem of interest and its needs/goals [51, 52]. By redistributing the weights of the sub-dimensions, it is also possible to trace back to approaches that focus more on certain aspects, such as task performance or safety, but which may be more "short-sighted" in the holistic assessment of collaboration. In Appendix A, more details were provided on the Simos' Procedure and its application to the framework. Once importance weights have been assigned to the sub-dimensions of the framework, the ELECTRE-II method can be used to generate a ranking of the different HRC configurations considered. Appendix B shows and explains in detail the various steps of the procedure for implementing the ELECTRE-II method. It is important to note that the application of this method can be easily automated, requiring only the dimension weights and evaluation profiles of the different HRC configurations as input. In order to validate and verify the robustness of the results obtained, it is possible to conduct a sensitivity analysis. This procedure consists in introducing random variations in the weights of the dimensions and verify possible changes in the ranking generated by ELECTRE-II. The smaller the deviations of the new rankings, the greater the robustness of the result originally obtained.

In order to show the implementation and practicability of the proposed methodology, a case study related to a collaborative parking-pawl assembly task in the automotive field was examined. Four alternative solutions of the collaborative task with comparable implementation and maintenance costs were presented and evaluated using the HRC framework. To identify the most preferable solution from a collaborative perspective, the four generated evaluation profiles were compared. As is often the case, an initial comparison did not reveal a profile that was clearly better than the others, highlighting the need to use a decision support method such as the one proposed. As far as the distribution of the weights of the HRC dimensions is concerned, it emerged that the team of experts gave particular importance to safety, human factors and ethical aspects, highlighting a humanoriented design strategy. From the application of the ELECTRE-II method, the HRC4 configuration was found to be the most preferable. The implementation of different communication means helped to increase the flexibility of interaction, allowing the operator to communicate with the robot in the mean that best suits the situation. In addition, the robot's ability to adapt to the state of the task helped make collaboration more natural. HRC1 ranked second, outperforming HRC3 and HRC2. Although HRC3 implements a natural communication mean (i.e., vocal commands), it increased the operator's workload significantly. HRC2 ranked last, mainly due to the relatively low performance in terms of *Trust* and *Safety*.

From an economic point of view, the considered HRC solutions have comparable implementation costs. Although the evaluation framework does not take the economic aspect directly into account, as it is focused on the evaluation of the quality of collaboration [21], it is possible to take it into account indirectly. In the design strategy definition phase, economic aspects can be indirectly quantified through the assignment of weights to individual sub-dimensions of the framework by the designers. Future work will focus on improving the framework through the explicit introduction of economic aspects.

7. CONCLUSIONS

The methodology proposed in this paper, combined with the evaluation framework of Gervasi *et al.* [21], allows to guide a team of experts in choosing between different solutions of the same collaborative task. One of the main advantages of this methodology is to embrace a holistic view of the problem, evaluating various aspects of the collaboration quality that take into account both technical and human aspects.

The automotive case study presented highlighted the potential difficulties in choosing between different design solutions considering all the multiple aspects of HRC. When comparing evaluations of different HRC solutions, it can often happen that one does not appear clearly better than the others. Furthermore, the management of evaluations of rather heterogeneous dimensions is not straightforward. The application of the proposed methodology allowed the group of experts to overcome these difficulties by: (i) formalizing a design strategy through the assignment of importance weights to the various (sub-)dimensions of the framework; (ii) generating a ranking of the solutions under consideration; (iii) verifying the robustness of the result obtained.

Future work will focus on improving the framework proposed by Gervasi *et al.* [21], such as (i) including more explicitly the economic aspects related to HRC implementation, (ii) identifying and deal with possible correlations between the sub-dimensions, and (iii) building benchmark profiles that support the construction of a unidimensional HRC scale.

AKNOWLEDGEMENTS

This work has been partially supported by "Ministero dell'Istruzione, dell'Università e della Ricerca" Award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6".

REFERENCES

1. Mateus, J.C., Claeys, D., Limère, V., Cottyn, J., Aghezzaf, E.-H.: A structured methodology for the design of a human-robot collaborative assembly workplace. Int. J. Adv. Manuf. Technol. 102, 2663–2681 (2019). https://doi.org/10.1007/s00170-019-03356-3
2. Pine, B.J.: Mass customization. Harvard business school press Boston (1993)
3. Cohen, Y., Shoval, S., Faccio, M.: Strategic View on Cobot Deployment in Assembly 4.0 Systems. IFAC-Pap. 52, 1519–1524 (2019). https://doi.org/10.1016/j.ifacol.2019.11.415
4. Huang, S., Ishikawa, M., Yamakawa, Y.: A coarse-to-fine framework for accurate positioning under uncertainties—from autonomous robot to human–robot system. Int. J. Adv. Manuf. Technol. 108, 2929–2944 (2020). https://doi.org/10.1007/s00170-020-05376-w
5. Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X.V., Makris, S., Chryssolouris, G.: Symbiotic human-robot collaborative assembly. CIRP Ann. 68, 701–726 (2019). https://doi.org/10.1016/j.cirp.2019.05.002
6. ISO/TS 15066:2016: Robots and robotic devices – Collaborative robots. International Organization for Standardization, Geneva, CH (2016)
7. Robla-Gómez, S., Becerra, V.M., Llata, J.R., González-Sarabia, E., Torre-Ferrero, C., Pérez-Oria, J.: Working Together: A Review on Safe Human-Robot Collaboration in Industrial Environments. IEEE Access. 5, 26754–26773 (2017). https://doi.org/10.1109/ACCESS.2017.2773127
8. Vicentini, F., Askarpour, M., Rossi, M.G., Mandrioli, D.: Safety Assessment of Collaborative Robotics Through Automated Formal Verification. IEEE Trans. Robot. 36, 42–61 (2020). https://doi.org/10.1109/TRO.2019.2937471
9. Gopinath, V., Johansen, K.: Understanding situational and mode awareness for safe human-robot collaboration: case studies on assembly applications. Prod. Eng. 13, 1–9 (2019). https://doi.org/10.1007/s11740-018-0868-2
10. Yoo, I.S., Landgraf, M., Ramer, C., Reitelshöfer, S., Ziegler, C., Franke, J.: My new colleague has artificial muscles: a DEA based approach for inherently compliant robotic systems. Prod. Eng. 8, 711–717 (2014). https://doi.org/10.1007/s11740-014-0564-9
11. Argall, B.D., Chernova, S., Veloso, M., Browning, B.: A survey of robot learning from demonstration. Robot. Auton. Syst. 57, 469–483 (2009). https://doi.org/10.1016/j.robot.2008.10.024
12. Raatz, A., Blankemeyer, S., Recker, T., Pischke, D., Nyhuis, P.: Task scheduling method for HRC workplaces based on capabilities and execution time assumptions for robots. CIRP Ann. 69, 13–16 (2020). https://doi.org/10.1016/j.cirp.2020.04.030
13. Faber, M., Mertens, A., Schlick, C.M.: Cognition-enhanced assembly sequence planning for ergonomic and productive human–robot collaboration in self-optimizing assembly cells. Prod. Eng. 11, 145–154 (2017). https://doi.org/10.1007/s11740-017-0732-9
14. Salm-Hoogstraeten, S. von, Müsseler, J.: Human Cognition in Interaction With Robots: Taking the Robot's Perspective Into Account: Hum. Factors. (2020). https://doi.org/10.1177/0018720820933764
15. Gervasi, R., Mastrogiacomo, L., Franceschini, F.: Towards the definition of a Human-Robot collaboration scale. In: Bini, M., Amenta, P., D'Ambra, A., and Camminatiello, I. (eds.)

Statistical Methods for Service Quality Evaluation - Book of short papers of IES 2019, Rome, Italy, July 4-5. pp. 75–80. Cuzzolin, Italy (2019)
16. Goodrich, M.A., Schultz, A.C.: Human-robot interaction: a survey. Now, Boston, Mass. (2007)
17. Beer, J.M., Fisk, A.D., Rogers, W.A.: Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction. J. Hum.-Robot Interact. 3, 74–99 (2014). https://doi.org/10.5898/JHRI.3.2.Beer
18. Bröhl, C., Nelles, J., Brandl, C., Mertens, A., Schlick, C.M.: TAM Reloaded: A Technology Acceptance Model for Human-Robot Cooperation in Production Systems. In: Stephanidis, C. (ed.) HCI International 2016 – Posters' Extended Abstracts. pp. 97–103. Springer International Publishing, Cham (2016)
19. Gualtieri, L., Palomba, I., Merati, F.A., Rauch, E., Vidoni, R.: Design of Human-Centered Collaborative Assembly Workstations for the Improvement of Operators' Physical Ergonomics and Production Efficiency: A Case Study. Sustainability. 12, 3606 (2020). https://doi.org/10.3390/su12093606
20. Rabbani, M., Behbahan, S.Z.B., Farrokhi-Asl, H.: The Collaboration of Human-Robot in Mixed-Model Four-Sided Assembly Line Balancing Problem. J. Intell. Robot. Syst. (2020). https://doi.org/10.1007/s10846-020-01177-1
21. Gervasi, R., Mastrogiacomo, L., Franceschini, F.: A conceptual framework to evaluate human-robot collaboration. Int. J. Adv. Manuf. Technol. 108, 841–865 (2020). https://doi.org/10.1007/s00170-020-05363-1
22. Gervasi, R., Digiaro, F., Mastrogiacomo, L., Maisano, D., Franceschini, F.: Comparing quality profiles in Human-Robot Collaboration: empirical evidence in the automotive sector. In: Proceedings book of the 4th International Conference on Quality Engineering and Management. pp. 89–114. , University of Minho, Portugal (2020)
23. Colgate, J.E., Wannasuphoprasit, W., Peshkin, M.A.: Cobots: robots for collaboration with human operators. In: Proceedings of the 1996 ASME International Mechanical Engineering Congress and Exposition. pp. 433–439 (1996)
24. Price, H.E.: The Allocation of Functions in Systems. Hum. Factors. 27, 33–45 (1985). https://doi.org/10.1177/001872088502700104
25. Charalambous, G., Fletcher, S., Webb, P.: Identifying the key organisational human factors for introducing human-robot collaboration in industry: an exploratory study. Int. J. Adv. Manuf. Technol. 81, 2143–2155 (2015). https://doi.org/10.1007/s00170-015-7335-4
26. Tan, J.T.C., Duan, F., Zhang, Y., Watanabe, K., Kato, R., Arai, T.: Human-robot collaboration in cellular manufacturing: Design and development. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 29–34. IEEE, St. Louis, MO, USA (2009)
27. Arai, T., Kato, R., Fujita, M.: Assessment of operator stress induced by robot collaboration in assembly. CIRP Ann. 59, 5–8 (2010). https://doi.org/10.1016/j.cirp.2010.03.043
28. Nelles, J., Kohns, S., Spies, J., Brandl, C., Mertens, A., Schlick, C.M.: Analysis of Stress and Strain in Head Based Control of Collaborative Robots—A Literature Review. In: Goonetilleke, R. and Karwowski, W. (eds.) Advances in Physical Ergonomics and Human Factors. pp. 727–737. Springer International Publishing, Cham (2016)
29. Eimontaite, I., Gwilt, I., Cameron, D., Aitken, J.M., Rolph, J., Mokaram, S., Law, J.: Language-free graphical signage improves human performance and reduces anxiety when working collaboratively with robots. Int. J. Adv. Manuf. Technol. 100, 55–73 (2019). https://doi.org/10.1007/s00170-018-2625-2
30. Sauppé, A., Mutlu, B.: The Social Impact of a Robot Co-Worker in Industrial Settings. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15. pp. 3613–3622. ACM Press, Seoul, Republic of Korea (2015)
31. Hoffman, G., Breazeal, C.: Effects of anticipatory perceptual simulation on practiced human-robot tasks. Auton. Robots. 28, 403–423 (2010). https://doi.org/10.1007/s10514-009-9166-3

32. Rajendran, V., Carreno-Medrano, P., Fisher, W., Werner, A., Kulić, D.: A Framework for Human-Robot Interaction User Studies. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 6215–6222 (2020)
33. Kopp, T., Baumgartner, M., Kinkel, S.: Success factors for introducing industrial human-robot interaction in practice: an empirically driven framework. Int. J. Adv. Manuf. Technol. 112, 685–704 (2021). https://doi.org/10.1007/s00170-020-06398-0
34. Thrun, S.: Toward a Framework for Human-robot Interaction. Hum-Comput Interact. 19, 9–24 (2004). https://doi.org/10.1207/s15327051hci1901&2_2
35. ISO 26800:2011: Ergonomics - General approach, principles and concepts. International Organization for Standardization, Geneva, CH (2011)
36. BS 8611:2016: Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems. British Standards Institution, London, UK (2016)
37. NIST: Framework for improving critical infrastructure cybersecurity. National Institute of Standards and Technology, Gaithersburg, MD, USA (2018)
38. Krüger, M., Wiebel, C.B., Wersing, H.: From Tools Towards Cooperative Assistants. In: Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17. pp. 287–294. ACM Press, Bielefeld, Germany (2017)
39. ISO 10218-2:2011: Robots and robotic devices – Safety requirements for industrial robots – Part 2: Robot systems and integration. International Organization for Standardization, Geneva, CH (2011)
40. ISO/TR 14121-2:2012: Safety of machinery – Risk assessment – Part 2: Practical guidance and examples of methods. International Organization for Standardization, Geneva, CH (2012)
41. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Hancock, P.A. and Meshkati, N. (eds.) Advances in Psychology. pp. 139–183. North-Holland (1988)
42. Yanco, H.A., Drury, J.: Classifying human-robot interaction: an updated taxonomy. In: 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583). pp. 2841–2846 vol.3 (2004)
43. Schaub, K., Caragnano, G., Britzke, B., Bruder, R.: The European Assembly Worksheet. Theor. Issues Ergon. Sci. 14, 616–639 (2013). https://doi.org/10.1080/1463922X.2012.678283
44. Bangor, A., Kortum, P.T., Miller, J.T.: An Empirical Evaluation of the System Usability Scale. Int. J. Human–Computer Interact. 24, 574–594 (2008). https://doi.org/10.1080/10447310802205776
45. Brooke, J.: SUS - A quick and dirty usability scale. In: Jordan, P., Thomas, B., Weerdmeester, B., and McClelland, I. (eds.) Usability Evaluation In Industry. pp. 189–194. CRC Press, London (1996)
46. Dedeke, A.: Cybersecurity Framework Adoption: Using Capability Levels for Implementation Tiers and Profiles. IEEE Secur. Priv. 15, 47–54 (2017). https://doi.org/10.1109/MSP.2017.3681063
47. Universal Robots: Collaborative robotic automation | Cobots from Universal Robots, https://www.universal-robots.com/
48. Figueira, J., Mousseau, V., Roy, B.: Electre Methods. In: Figueira, J., Greco, S., and Ehrogott, M. (eds.) Multiple Criteria Decision Analysis: State of the Art Surveys. pp. 133–153. Springer, New York, NY (2005)
49. Maystre, L.Y., Pictet, J., Simos, J.: Méthodes multicritères ELECTRE: description, conseils pratiques et cas d'application à la gestion environnementale. PPUR presses polytechniques (1994)
50. Figueira, J., Roy, B.: Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. Eur. J. Oper. Res. 139, 317–326 (2002). https://doi.org/10.1016/S0377-2217(01)00370-8

51. Franceschini, F., Maisano, D.: Fusing incomplete preference rankings in design for manufacturing applications through theZMII-technique. Int. J. Adv. Manuf. Technol. 103, 3307–3322 (2019). https://doi.org/10.1007/s00170-019-03675-5
52. Franceschini, F., Maisano, D.: Aggregation of incomplete preference rankings: Robustness analysis of the ZMII-technique. J. Multi-Criteria Decis. Anal. 27, 337–356 (2020). https://doi.org/10.1002/mcda.1721
53. Ostanello, A.: Outranking Methods. In: Fandel, G. and Spronk, J. (eds.) Multiple Criteria Decision Methods and Applications: Selected Readings of the First International Summer School Acireale, Sicily, September 1983. pp. 41–60. Springer, Berlin, Heidelberg (1985)
54. Godsil, C., Royle, G.F.: Algebraic Graph Theory. Springer-Verlag, New York (2001)

APPENDIX A – DESCRIPTION OF SIMOS' PROCEDURE APPLICATION

This appendix provides a brief description of Simos' procedure and its application to the problem of interest for assigning weights to the HRC (sub-)dimensions.

The Simos' procedure is an indirect method for assigning weights to criteria that has been widely used in decision making problems, as it represents a relatively easy method for decision-makers to express their preferences [49, 50]. The selection of weights is performed by asking decision-makers to express the relative importance of the criteria through the arrangement of criteria cards (i.e., a set of cards containing the name of each criterion), from the least to the most important one. If some criteria have the same importance for decision-makers, they can be placed together in the same rank. As a result of this arrangement, a complete pre-order of all criteria is generated. Decision-makers also have a set of blank cards, which can be inserted between two successive criteria (or two successive subsets of *ex aequo* criteria). These white cards allow to increase the difference of importance of successive criteria: the more the number of white cards between two successive criteria, the greater the difference between their importance. The non-normalized weight of each rank is derived by dividing the sum of card positions by the number of cards (see Table A.1 and A.2). Afterwards, the normalized weight of each rank is obtained by dividing the non-normalized weight by the sum of the card positions of the criteria (without considering white cards).

In the problem of interest, the Simos' procedure has been applied to both the main HRC dimensions and their respective sub-dimensions. Moreover, no white cards were introduced by the team of expert. Table A.1 contains the results of the Simos' procedure applied by the team of expert to the main HRC dimensions. Within each dimension, the team of experts used the Simos' procedure to assign a relative weight to the sub-dimensions, as shown in Table A.2. The final weights of the (sub-)dimensions, used as input to the ELECTRE-II method, were derived by the product between the normalized relative weights and the normalized weight of their respective main HRC dimension (Table A.2).

Table A.1 – Assigning weights to HRC dimensions by using Simos' procedure.

Rank	Subsets of *ex aequo*	Number of cards	Positions	Non-normalized weight	Normalized weight
1	{Cybersecurity}	1	1	1	$\frac{1}{28} \cong 0.04$
2	{Autonomy, Information Exchange}	2	2, 3	$\frac{2+3}{2} = 2.5$	$\frac{2.5}{28} \cong 0.09$
3	{Adaptivity and Training}	1	4	4	$\frac{4}{28} \cong 0.14$
4	{Task, Ethics}	2	5, 6	$\frac{5+6}{2} = 5.5$	$\frac{5.5}{28} \cong 0.20$
5	{Human Factors}	1	7	7	$\frac{7}{28} = 0.25$

Table A.2 – Assigning weights to sub-dimensions of each HRC dimension by using Simos' procedure.

Rank	Subsets of *ex aequo*	Number of cards	Positions	Non-normalized weight	Normalized weight	Final weight
			Information Exchange (9%)			
1	{Communication medium}	1	1	1	$\frac{1}{3} \cong 0.33$	2.97%
2	{Communication format}	1	2	2	$\frac{2}{3} \cong 0.67$	6.03%
			Adaptivity and Training (14%)			
1	{Robot training method}	1	1	1	$\frac{1}{6} \cong 0.17$	2.38%
2	{Operator training}	1	2	2	$\frac{2}{6} \cong 0.33$	4.62%
3	{Robot adaptivity}	1	3	3	$\frac{3}{6} = 0.50$	7.00%
			Task (20%)			
1	{Performance}	1	1	1	$\frac{1}{3} \cong 0.33$	6.60%
2	{Safety}	1	2	2	$\frac{2}{3} \cong 0.67$	13.40%
			Human Factors (25%)			
1	{Trust, Usability}	2	1, 2	$\frac{1+2}{2} = 1.5$	$\frac{1.5}{10} = 0.15$	3.75%
2	{Workload, Physical Ergonomics}	2	3, 4	$\frac{3+4}{2} = 3.5$	$\frac{3.5}{10} = 0.35$	8.75%
			Ethics (20%)			
1	{Social impact, Social acceptance}	2	1, 2	$\frac{1+2}{2} = 1.5$	$\frac{1.5}{3} = 0.50$	10.00%
			Cybersecurity (4%)			
1	{Identification, Protection, Detection, Response, Recovery}	5	1, 2, 3, 4, 5	$\frac{1+2+3+4+5}{5} = 3$	$\frac{3}{15} = 0.20$	0.80%

APPENDIX B – DESCRIPTION OF THE ELECTRE-II APPLICATION

This appendix provides a simplified description of ELECTRE-II [48] and its application to the problem of interest.

Initially, (sub-)dimensions should be hierarchized through (cardinal) weights. Table B.1 contains the resulting weight for each HRC sub-dimension assigned by the team of expert (see Appendix A); in addition, it includes the so-called *profiles*, depicting the performance of the alternative HRC configurations (a_1 to a_4)[1], from the viewpoint of each HRC sub-dimension.

In the initial phase of ELECTRE-II, the profiles (a_1 to a_4) are turned into sets of paired-comparison relationships of strict preference ("≻" or "≺") or indifference ("~"), and then aggregated into a single set of *outranking* relationships. The total number of paired comparisons is $2 \cdot \binom{n}{2} = n \cdot (n-1)$, as it includes both "direct" and "reverse" comparisons, e.g., (a_1, a_2) and (a_2, a_1). Since the problem of interest includes n = 4 profiles, there are twelve total paired comparisons (see Table B.2).

This phase of ELECTRE-II can be decomposed in three steps:

1. For each paired comparison (a_i, a_j), the sub-dimensions are grouped in the three sets J^+, $J^=$ and J^-, for which $a_i \succ a_j$, $a_i \sim a_j$, and $a_i \prec a_j$, respectively. Table B.2 contains the construction of the three sets for the sub-dimensions of the problem of interest and the respective weights.

2. The three sets J^+, $J^=$ and J^- are associated with some so-called *consistency scores* W^+, $W^=$ and W^-, i.e., the sums of the w_j values related to the sub-dimensions (e_k) contained in them. Consistency scores for the paired comparison (a_i, a_j) are formally defined as:

$$
\begin{aligned}
W^+ &= \sum_{k \in J^+} w_k, \text{ where } J^+ = \left\{ e_k : g_{e_k}(a_i) > g_{e_k}(a_j) \right\} \\
W^= &= \sum_{k \in J^=} w_k, \text{ where } J^= = \left\{ e_k : g_{e_k}(a_i) = g_{e_k}(a_j) \right\} \\
W^- &= \sum_{k \in J^-} w_k, \text{ where } J^- = \left\{ e_k : g_{e_k}(a_i) < g_{e_k}(a_j) \right\}
\end{aligned}
\tag{B.1}
$$

   where $g_k(\cdot)$ is a function that returns the scale value for the sub-dimension k.

   Table B.3 contains the calculation of these scores with reference to the J^+, $J^=$, and J^- sets in Table B.2.

[1] The more compact notation "a_1, a_2, …" will be adopted to indicate the corresponding HRC configurations "HRC_1, HRC_2, …" described in Sect. 3.

Table B.1 – List of the weights related to each sub-dimension, with relevant performance profiles.

Sub-dimension	Label	Weight (%)	Profiles			
			a_1	a_2	a_3	a_4
Autonomy	e_1	9.00	L3	L3	L3	L3
Communication medium	e_2	2.97	L2	L2	L3	L3
Communication format	e_3	6.03	L1	L1	L2	L3
Robot adaptivity	e_4	7.00	L1	L1	L1	L2
Robot training method	e_5	2.38	L1	L1	L1	L2
Operator training	e_6	4.62	L2	L2	L2	L1
Performance	e_7	6.60	L2	L3	L2	L3
Safety	e_8	13.40	L3	L2	L3	L3
Workload	e_9	8.75	L2	L2	L1	L2
Trust	e_{10}	3.75	L3	L2	L3	L3
Physical ergonomics	e_{11}	8.75	L2	L2	L2	L2
Usability	e_{12}	3.75	L1	L1	L1	L2
Social impact	e_{13}	10.00	L1	L1	L1	L1
Social acceptance	e_{14}	10.00	L2	L2	L2	L1
Identification	e_{15}	0.80	L1	L1	L1	L2
Protection	e_{16}	0.80	L1	L1	L1	L2
Detection	e_{17}	0.80	L1	L1	L1	L2
Response	e_{18}	0.80	L1	L1	L1	L2
Recovery	e_{19}	0.80	L1	L1	L1	L2

Table B.2 – Sets J^+, $J^=$, and J^- of the problem of interest in the initial phase of the ELECTRE-II method.

Paired comparison		Sets of sub-dimensions		
		J^+	$J^=$	J^-
1	a_1, a_2	$\{e_8, e_{10}\}$	$\{e_1, e_2, e_3, e_4, e_5, e_6, e_9, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_7\}$
2	a_1, a_3	$\{e_9\}$	$\{e_1, e_2, e_4, e_5, e_6, e_7, e_8, e_{10}, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_3\}$
3	a_1, a_4	$\{e_6, e_{14}\}$	$\{e_1, e_8, e_9, e_{10}, e_{11}, e_{13}\}$	$\{e_2, e_3, e_4, e_5, e_7, e_{12}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$
4	a_2, a_1	$\{e_7\}$	$\{e_1, e_2, e_3, e_4, e_5, e_6, e_9, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_8, e_{10}\}$
5	a_2, a_3	$\{e_7, e_9\}$	$\{e_1, e_2, e_4, e_5, e_6, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_3, e_8, e_{10}\}$
6	a_2, a_4	$\{e_6, e_{14}\}$	$\{e_1, e_7, e_9, e_{11}, e_{13}\}$	$\{e_2, e_3, e_4, e_5, e_8, e_{10}, e_{12}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$
7	a_3, a_1	$\{e_3\}$	$\{e_1, e_2, e_4, e_5, e_6, e_7, e_8, e_{10}, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_9\}$
8	a_3, a_2	$\{e_3, e_8, e_{10}\}$	$\{e_1, e_2, e_4, e_5, e_6, e_{11}, e_{12}, e_{13}, e_{14}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_7, e_9\}$
9	a_3, a_4	$\{e_6, e_{14}\}$	$\{e_1, e_8, e_{10}, e_{11}, e_{13}\}$	$\{e_2, e_3, e_4, e_5, e_7, e_9, e_{12}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$
10	a_4, a_1	$\{e_2, e_3, e_4, e_5, e_7, e_{12}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_1, e_8, e_9, e_{10}, e_{11}, e_{13}\}$	$\{e_6, e_{14}\}$
11	a_4, a_2	$\{e_2, e_3, e_4, e_5, e_8, e_{10}, e_{12}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_1, e_7, e_9, e_{11}, e_{13}\}$	$\{e_6, e_{14}\}$
12	a_4, a_3	$\{e_2, e_3, e_4, e_5, e_7, e_9, e_{12}, e_{15}, e_{16}, e_{17}, e_{18}, e_{19}\}$	$\{e_1, e_8, e_{10}, e_{11}, e_{13}\}$	$\{e_6, e_{14}\}$

3. These indicators are then used to perform two tests, which are based on the two major concepts of: (i) *concordance*, i.e., for the first configuration (a_i) to outperform the second one (a_j), a sufficient majority of sub-dimensions should be in favor of this assertion;
(ii) *non-discordance*, i.e., when the concordance condition holds, none of the sub-dimensions in the minority (i.e., those $\in J^-$) should oppose too strongly to the assertion.

By translating these concepts into operational terms, two tests are defined [48, 53]. The first one is called *concordance test*, which is based on the verification of both the conditions:

$$\begin{matrix} \dfrac{W^+ + W^=}{W} \geq c_1 \\ \dfrac{W^+}{W^-} \geq c_2 \end{matrix}, \tag{B.2}$$

being $W = W^+ + W^= + W^-$, and $c_1 = 0.7$ and $c_2 = 1$ two conventional thresholds.
The second test is called *non-discordance* test, which verifies the absence of a so-called *veto* on the outranking of a_j by a_i (i.e., a_iSa_j, where the symbol "S" denotes the outranking relationship) on sub-dimensions belonging to the J^- set. The veto may occur for sub-dimensions for which the scale value of the j-th configuration is much higher than that of the i-th configuration, i.e., $g_k(a_j) \succ\succ g_k(a_i)$. More precisely, a veto occurs for a generic sub-dimension k expressed on a cardinal scale, if:

$$g_k(a_j) - g_k(a_i) > v_k \tag{B.3}$$

In the problem addressed, since (i) sub-dimensions are expressed on ordinal scales and (ii) there is no particular reason for the veto condition to be more stringent for some sub-dimensions than for others, the same veto threshold, corresponding to a difference of two levels of the scale in use, has been set for all the sub-dimensions (i.e., $v_k = v = 2 \ \forall k \in \{e_1, \ldots, e_{19}\}$). Therefore, referring to the totality of the sub-dimensions possibly included in the J^- set, the most unfavourable condition to verify if a veto on the a_iSa_j outranking condition does not occur (i.e., if the non-discordance test is passed) is:

$$\max_{k \in J^-} \left[g_k(a_j) - g_k(a_i) \right] \leq v \tag{B.4}$$

When, for a generic paired comparison (a_i, a_j), the concordance and non-discordance test are both passed, it can be stated that a_iSa_j.
In the problem addressed, the paired comparisons (a_1, a_2), (a_1, a_3), (a_3, a_2), (a_4, a_1), (a_4, a_2), and (a_4, a_3) passed both the concordance and non-discordance tests (see Table B.3). Next, results can be visualized in a graph, in which (i) vertices represent configurations and (ii) edges joining two vertices represent the corresponding outranking relationship (Figures B.1 and B.2).

Table B.3 - Determination of consistency scores and outranking relationships (a_iSa_j), in the initial phase of the ELECTRE-II method, with reference to the J^+, $J^=$ and J^- sets in Table B.2.

Paired comparison		Consistency scores			Concordance test				Non-discordance test		a_iSa_j?
		W^+	$W^=$	W^-	$(W^+ + W^=)/W$ ($\geq c_1$)		W^+/W^- ($\geq c_2$)		$\max_{k \in J^-}[g_k(a_j) - g_k(a_i)] \leq v$		
1	a_1, a_2	0.172	0.773	0.066	0.944	✓	2.598	✓	1	✓	Yes
2	a_1, a_3	0.088	0.862	0.060	0.950	✓	1.451	✓	1	✓	Yes
3	a_1, a_4	0.146	0.537	0.327	0.683	×	0.447	×	N/A	N/A	No
4	a_2, a_1	0.066	0.773	0.172	0.839	✓	0.385	×	N/A	N/A	No
5	a_2, a_3	0.154	0.625	0.232	0.778	✓	0.662	×	N/A	N/A	No
6	a_2, a_4	0.146	0.431	0.433	0.577	×	0.338	×	N/A	N/A	No
7	a_3, a_1	0.060	0.862	0.088	0.923	✓	0.689	×	N/A	N/A	No
8	a_3, a_2	0.232	0.625	0.154	0.857	✓	1.510	✓	1	✓	Yes
9	a_3, a_4	0.146	0.449	0.415	0.595	×	0.352	×	N/A	N/A	No
10	a_4, a_1	0.327	0.537	0.146	0.864	✓	2.239	✓	1	✓	Yes
11	a_4, a_2	0.433	0.431	0.146	0.864	✓	2.960	✓	1	✓	Yes
12	a_4, a_3	0.415	0.449	0.146	0.864	✓	2.837	✓	1	✓	Yes

Thresholds are set as: $c_1 = 0.7$, $c_2 = 1$ and $v_k = v = 2$ $\forall k \in \{e_1, \ldots, e_{19}\}$.
The symbols ✓ and × denote whether a certain test (or part of it) is passed or not.

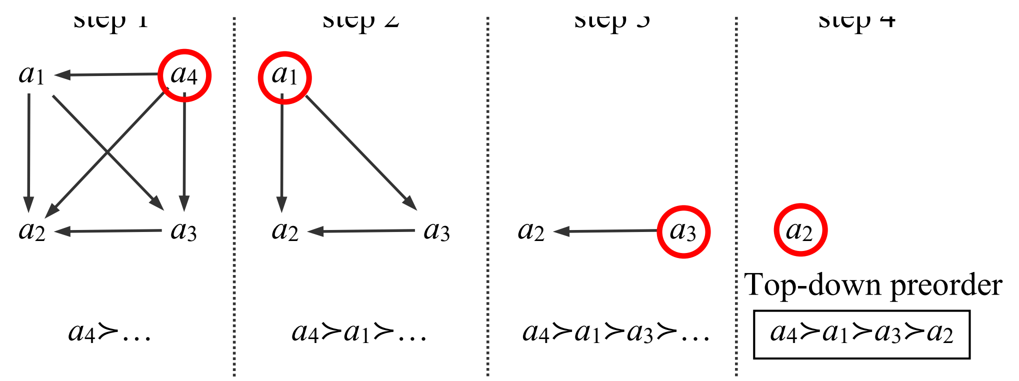


Figure B.1 - Basic steps for transforming the outranking graph (Step 1) in a top-down pre-order (Step 4), in the second phase of the ELECTRE-II method; the configurations selected in each step are circled.

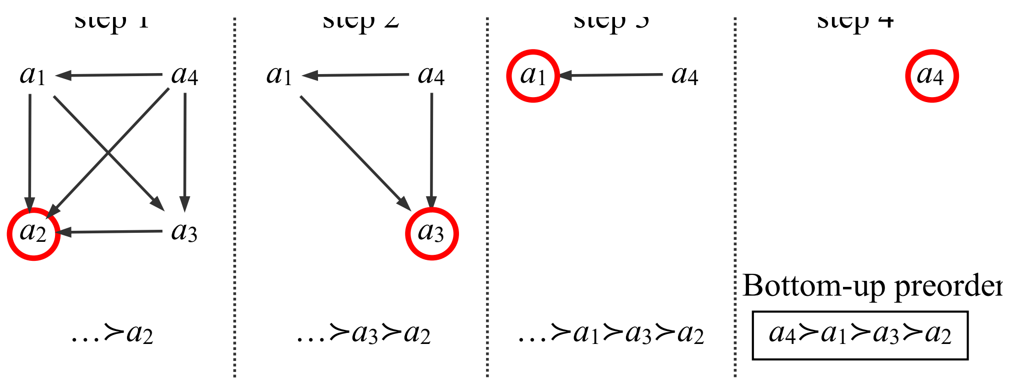


Figure B.2 - Basic steps for transforming the outranking graph (Step 1) in a bottom-up pre-order (Step 4), in the second phase of the ELECTRE-II method; the configurations selected in each step are circled.

The second phase of the ELECTRE-II method is aimed at deriving a collective ranking from the outranking relationships. To this purpose, the outranking relationships are turned into so-called *pre-orders*, either using a *top-down* and *bottom-up* procedure [48, 53].

For both the preorders, the preliminary step is to identify and eliminate possible circuits, i.e., those combinations of outranking relationships in which the transitivity property is violated (e.g., a_1Sa_2, a_2Sa_3 and a_3Sa_1). For each of the circuits identified, the corresponding outranking relationships are deleted, then the configurations are grouped in the same class and considered as indifferent (e.g., $a_1 \sim a_2 \sim a_3$) [54]. In our particular case, the top-down and bottom-up preorders coincides (Figures B.1 and B.2), hence the collective ranking resulting from the application of the ELECTRE-II method is $a_4 \succ a_1 \succ a_3 \succ a_2$.