Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002

Design Principles for Safe Human Robot Collaboration

Laura Tomidei[1], Matthias Guertler[1,3], Nathalie Sick[1,3], Gavin Paul[2,3], Marc Carmichael[2,3]

[1] Centre for Advanced Manufacturing, University of Technology Sydney, Australia
[2] Robotics Institute, University of Technology Sydney, Australia
[3] Australian Cobotics Centre (ITTC for Collaborative Robotics in Advanced Manufacturing)

Abstract. With the development of collaborative robots (cobots), a paradigm shift in human-robot collaboration (HRC) is emerging in the workplace. When introducing cobots, a new range of hazards and harms needs to be considered. While physical hazards have been extensively studied and were paramount in the development of cobots, lesser-known hazards are related to mental and ethical wellbeing. Accordingly, most existing safety measures are designed to address exclusively physical hazards including ergonomics. To this end, this study sets out to develop holistic design principles for safe HRC by adopting a human-centred approach. A systematic review of the relevant literature combined with real-world insights gathered through interviews with industry and academic experts leads to design principles for safe HRC that can contribute to the future development of collaborative robot systems. This also highlights challenges which future research around safety guidelines and standards needs to address.

Keywords: human robot collaboration, cobot, collaborative robot, ergonomics, risk, hazard, harm, safety, safe design, design principles.

1 Introduction

Due to industry trends such as the shift from mass production to mass customisation, human robot collaboration (HRC) - the collaboration between human operators and robotic assistive systems - has become more attractive [1].

Collaborative robots or *cobots* represent a new type of robotic system capable of HRC that opens up possibilities for deploying robots in industrial workplaces. Even though such physical robotic assistance promises many advantages on the factory floor, the fusion of the humans' and robots' workspaces entails various risks to operators. In this context, safety measures are standard practice in traditional industrial settings, but these work environments are radically changing through the adoption of HRC.

International standards specify safety measures for robots, robotic systems, and their integration [2, 3]. Despite the formal coverage in standards, the current strategies to ensure safe HRC appear limited to mitigating physical harm and working environments that produce industrial and manufacturing goods. However, advanced systems that allow human-robot interaction have the potential to serve both industrial and service-oriented domains, with possible future applications ranging from industrial co-workers, to mobile servants over robots in the professional service sector [4]. As a result, there is a growing need for safety standards and practices that address safe HRC in a wider array of work environments, cobot applications, and organisations that will use cobots.

Compared to traditional robots, cobots are designed to operate in close proximity to humans. They may introduce entirely different or additional hazards and harms, which range from increased hazardous contacts to psychological discomfort due to factors such as fear of job loss or loss of work agency. In practice, most safety measures are focussed on the physical hazards HRC may cause. However, the growing prevalence of cobots emphasizes the need to consider psychological and ethical dimensions linked to mental wellbeing rather than just physical safety. Thus far, there appears to be little consideration for how mental wellbeing can be affected by HRC. For example, anxiety and stress within workers can be caused not only by their work conditions near robots but also by job precarity and fear of losing their role [5]. For this reason, in the context of safe HRC, particular attention needs to be given to human factors, aiming to ensure both human physical and mental wellbeing. However, the variety of HRC systems and associated workplaces are a challenge to universal sets of HRC safety factors. This leads to the following research question: *What are the overarching design principles for safe HRC?* To answer this question, using a design-led approach, this study first provides a human-centred understanding of hazards that HRC may cause and their root causes, encompassing physical and mental wellbeing. Secondly, existing safety measures are identified and mapped against the hazards and root causes they allow to mitigate. Based on this analysis, design principles for safe HRC are developed.

The remainder of this article is structured as follows. The background section outlines the existing standards and definitions of HRC, while the methodology section presents the research design used for this study. The findings are structured into two sections. Firstly, HRC hazards and safety measures are presented. Existing safety measures are mapped against hazards and root causes to identify potential gaps. Second, the design principles derived from the mapping of hazard and safety measures are introduced. The conclusion section discusses the limitations and potential avenues for future research.

2 Background

While industrial robots have been around for decades, only in recent years have we seen the rise of collaborative robots. Unlike conventional industrial robots, cobots are designed to be operated in a shared workspace with humans. Using new methods such as lightweight construction, rounded or padded corners, inbuilt force and/or torque sensing, or mechanical compliance, cobots are able to be utilised without the need for safety cages or active safety devices, enabling HRC [6].

A collaborative workspace is defined as an "operating space where the robot system (including the workpiece) and a human can perform tasks concurrently during production operation" (see Fig. **1**) [2, 3, 7]. In this type of operation, "operators can work in close proximity to a robot system while power to the robot's actuators is available, and physical contact between an operator and the robot system can occur within a collaborative workspace" [7]. These technical specifications detail features and safety requirements for operating robots in different modes of collaboration. Robots designed for collaborative operation are required to comply with one or more of the following features: safety-rated monitored stop, hand guiding, speed and separation monitoring, and power and force limiting [3].

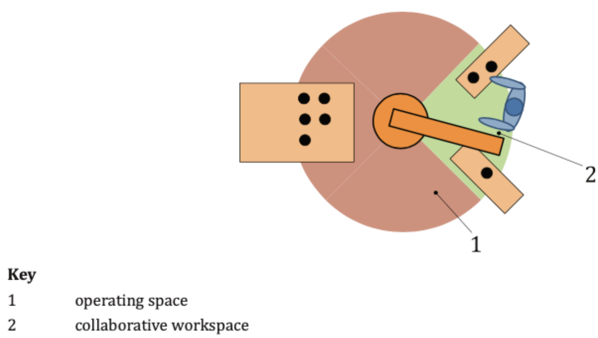


Fig. 1. Example of a collaborative workspace [7]

A robot can be considered as collaborative when; (a) it shares the workspace with a human, (b) tasks are performed at the same time and sometimes require physical contact, and (c) the robot's features include one or more of the four safety modes specified by the standards [2, 3].

Based on these definitions, it is clear that human-robot collaboration is a complex socio-technical system requiring the interaction between humans, machines, and other environmental aspects [8, 9]. For this reason, these systems need to be analysed using frameworks that consider all the relevant dimensions. When analysing HRC, the socio-technical context points to four main dimensions, namely

Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002

(1) human operator, (2) cobot, (3) working system, and (4) enterprise and contextual. Not only do the human operator and the cobot need to be considered individually, but also the working system or the cell design of the collaborative space where human and robot interact. In addition, enterprise and contextual factors play a role, such as task processes, roles or responsibilities, and workforce training [10] (see Fig. **2**). Existing standards describe basic hazards associated with robotic systems, while acknowledging that their key sources are frequently unique to the specific system [3]. To this end, the indications mainly refer to the robot and the robotic system, with only a few mentions to broader organisational factors such as different types of training. In addition to this, the scope is almost exclusively limited to the physical wellbeing of humans. Harms such as stress, fatigue and lack of concentration receive only limited coverage [7].

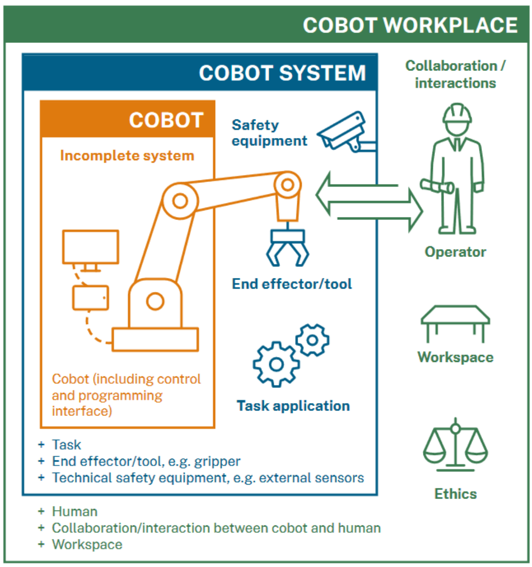


Fig. 2. System levels of a human-robot collaboration [11]

In practice, safety can only be achieved when HRC is able to protect both human physical and mental wellbeing. The consideration of the key dimensions of physical, cognitive, and organisational factors is crucial to represent the socio-technical complexity of HRC. Human safety is a key factor in the facilitation of human and robot coexistence [12]. When analysing the conditions that make HRC safe, this study adopts a human-centred perspective on safety-based ergonomics or human factors (HF/E) (see Fig. **3**). In fact, HF/E is defined as "*the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data, and methods to design in order to optimize human wellbeing and overall system* *performance"* [13, 14]. In other words, HF/E involves *"conducting research regarding human psychological, social, physical, and biological characteristics, maintaining the information obtained from that research, and working to apply that information with respect to the design, operation, or use of products or systems for optimizing human performance, health, safety, and/or habitability"* [15].

The key dimensions of HF/E are not only physical factors, typically related to physical activity, but also cognitive factors, concerned with metal processes and related aspects (e.g. mental workload, stress), and organisational factors, related to participation and collaboration in socio-technical systems [14]. The interplay of these three key dimensions is well-placed to represent the socio-technical complexity of HRC and highlights that safety is achieved only when HRC is able to protect both human physical and mental wellbeing.

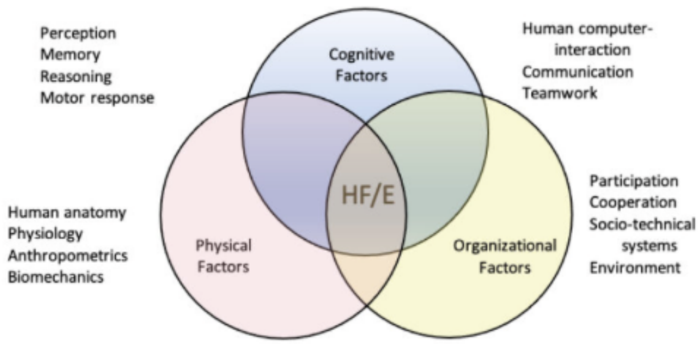


Fig. 3. Safety domains of ergonomics or human factors to address socio-technical complexity of HRC [14]

Thus, to ensure a safe design of HRC systems, it is crucial to understand potential hazards and what might cause them in the first place. Despite this, much is unknown about the hazards and harms of HRC, and even fewer research has been done on the possible root causes. While risks represent the possibility of negative events that can occur with a certain probability induced by unplanned circumstances, hazards are defined as *a thing that can be dangerous or cause damage* (to somebody/something) (Oxford dictionary). Accordingly, harms are *damages or injuries that are caused by a person or an event* (Oxford dictionary). Physical hazards such as the movement of any part of the cobot arm or end-effector, contact between fixtures, and failure of safeguarding devices have been extensively studied [2, 3] and were at the paramount in the development of HRC. Lesser known are other forms of harm such as mental strain and emotional stress due by the potential loss of work agency, fear of job loss, and fear of losing contact with colleagues [16]. This wide range of hazards is generated by the close interaction that HRC requires.

Understanding and categorising the key hazards and root causes considering physical, cognitive, and organisational factors is key for a human-centred design of HRC. At the same time, safety measures may be present at all levels of the collaborative system (i.e. cobot, cobot system, and workplace) and all aspects should be considered to ensure safety and address all potential hazards. By mapping and comparing the wide range of hazards against existing safety measures, it is possible to identify potential gaps and derive principles for safe and human-centred design of HRC.

3 Methodology

Due to the socio-technical and practice-related aspects that affect HRC, this study uses a design-led approach, comprising three steps, to holistically investigate technical as well as human-oriented perspectives. As a foundational step, a structured literature review of HRC safety builds the conceptual basis of relevant theoretical contributions [17]. Based on these insights, expert interviews are an effective methodology to complement and contribute to a conceptual body of knowledge by exploring meaning and perceptions among stakeholders and gaining a better understanding of phenomena [18]. In the third step, all findings are integrated into design principles for safe HRC.

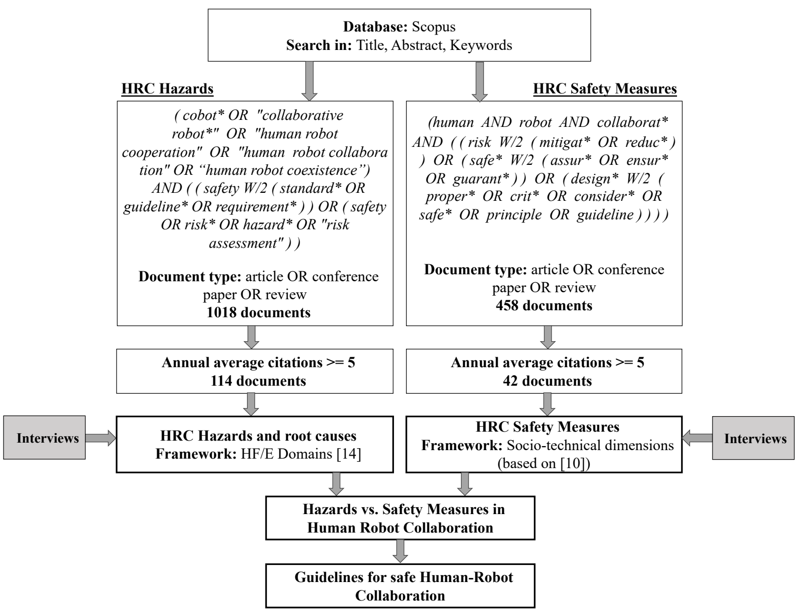


Fig. 4. Research Design of Data Collection and Analysis

Fig. **4** shows an overview of the research design, which is explained in the following.

3.1 Data Collection

3.1.1 Literature review. The structured literature review is based on Scopus, as it provides high-quality scholarly literature from a variety of scientific fields [19]. The literature search used two search strings, a hazard and a safety oriented one, that were defined in accordance with the research aims and scope, based on the definitions of cobots used in existing literature.

The first search string focused on HRC-related hazards and harms, using synonyms of cobots and synonyms of safety requirements or risks. The second search string aimed at HRC safety design, using synonyms of cobots combined with synonyms of risk mitigation and methodical design support. While the first search string used more specific cobot synonym terms to reduce the otherwise vast number of false-positive findings, the second search string required looser cobot terms to ensure enough findings (see

Fig. **4**). The documents identified by the two search strings have been analysed in parallel. The Scopus search returned a total of 1,018 hazard-related and 458 safety measure-related initial documents (see

Fig. **4**). From this dataset, only the most important and influential documents in the field with a minimum of 5 annual citations on average are included (n = 114 and n = 42). After this pre-filtering, the documents were screened concerning their relevance based on the following inclusion criteria (logical or):

- Describes hazards and harms of HRC,
- Describes approaches for HRC risk assessment,
- Focuses on analysing, listing, or mapping risks arising from HRC,
- Addresses design measures to support safety in HRC,
- Illustrates approaches to mitigate risks of HRC in a socio-technical system, or
- Focuses on HRC in an industrial environment (use case) or discusses design principles that can be transferred to HRC

The aim of the two independent literature reviews was twofold. On one hand the findings allowed to gain an understanding of the domain knowledge in the existing

DOI: 10.55612/s-5002-061-002

literature. However, to obtain a thorough understanding of industry practices, a supplementary interview study was required to support and strengthen the validity of the results. Thus, on the other hand, the literature reviews informed the design of the methodology for the interview study.

3.1.2 Interview study. The interview study employed a contextual inquiry approach [20]. The study explored the human attitudes and perceptions of various stakeholders across the cobot industry to understand how safe design can be supported and enabled. A combination of purposive sampling and snowball recruitment strategies was used to capture a wide array of perspectives on safe human-robot collaboration. The intentionally broad inclusion criteria (see Table 1) ensured that the study included a diverse set of participants, use-cases and industry sectors. Considering the unclear definition and use-cases between collaborative and industrial robots, the recruitment did not exclude research participants who had only experience working with industrial robots. This inclusion was especially important to recruit participants from non-traditional cobot user organisations, who do not have robotic, manufacturing, or engineering backgrounds and may not understand the nuanced differentiations between different robot types.

Initially, 70 individuals and organisations were identified as suitable research participants and contacted via recruitment phone calls and/or email requests. This resulted in a total of 19 interviews with 22 participants. Three of those were group interviews, each with two participants from the same organisation.

The semi-structured, one-hour interviews started after receiving the ethics approval and were conducted online using 'Microsoft Teams', following a semi-structured interview guide (UTS HREC REF NO. ETH21-6244). All interviews were recorded and uploaded to the online transcription and coding platform 'Condens'. The qualitative content analysis used an adapted version of the open, axial and selective coding approach borrowed from grounded theory [21]. Two members of the research team were tasked with reading and coding the transcription according to the hazard categories, risk mitigation strategies, and the socio-technical dimensions of design as identified by the literature review. New labels were created when the research team identified gaps in the existing label categories and literature review to highlight emerging patterns and themes. The label categories included; cobot definition, corporate environment, equipment selection, ethical, guideline recommendations, physical, process, psychosocial/ergonomic, role and responsibilities, task assignment, training, and workspace design.

Table 1. Participants to the interview study

No#	Interview participant category	Industry/sector	Participant position title
1	Cobot User	Tertiary Education	Coordinator/ Technician
2	Potential Cobot User	Food	Operational Manager
3	Distributor, Supplier, Integrator	Robotics/ Automation	Electronic Engineer

4	Distributor, Supplier, Integrator	Robotics / Automation	Founder & Project Manager
5	Industry Partner (*Risk Assessor*)	Independent Product Safety Assessors	Director
6	Industry Partner (Risk Assessor)	Independent Product Safety Assessors	Business Development Manager
7	Cobot User	Film	Director of Photography & Senior Motion Control Operator
8	Manufacturer	Safety Peripheral Equipment	Chief Technology Officer
9	Industry Partner (Risk Assessor)	Work, Health, and Safety	Work Health and Safety Inspector
10	Industry Partner (Researcher)	Robotics	Professor
11	Industry Partner (Researcher)	Robotics	Senior Lecturer
12	Supplier	Robotics/ Automation	Business Development Manager
13	Integrator	Robotics/ Automation	Director
14	Integrator + Cobot User	Higher Education	CEO
15	Supplier + Integrator	Robotics/ Automation	Project Engineer
16	Cobot User	Physical Rehabilitation	CEO & Founder
17	Cobot User	Custom Manufacturing	Operator and Head of Finishing
18	Industry Partner (Researcher)	Advanced Manufacturing	Professor & Centre Director
19	Cobot User	Higher Education	Technical Officer
20	Cobot User	Custom Manufacturing	Operational Manager
21	Supplier	Cobot Manufacturer	Operational Manager
22	Supplier	Cobot Manufacturer	Sales Engineer

3.2 Data Analysis and Synthesis

Together, the results from the two independent literature reviews and the interview study were analysed using the frameworks presented in section 2. Hazards findings were categorised using the cause-effect analysis and associated to the HF/E factors (see Fig. **3**). Findings related to the HRC safety measures were clustered using System levels of a human-robot collaboration [10]: (a) humans to be kept safe, (b) cobot, (c) working system, and (d) enterprise and context (see Fig. **2**). Comparing these two sets of results allowed to highlight addressed hazards and existing gaps of safety measures as well as to derive the overarching design principles for safe HRC.

To derive these design principles, three interdisciplinary researchers of the team (mechanical engineer, social designer, and work health and safety expert) used a qualitative content analysis approach [21]. Independently, they analysed the identified safety measures concerning underlying themes, such as the need to protect cobots and safety measures against tampering. The individual themes were discussed and consolidated into five themes, representing five design principles of safe HRC. Subsequently, the other eight members of the research team reviewed the principles and provided feedback to evaluate and fine-tune the principles (see Fig. 5).

4 Results

4.1 Human-Robot Collaboration Hazards and Safety Measures

The analysis of HRC hazards and their root causes demonstrates that safety is achieved not only by considering physical factors, but also cognitive and organisational ones, despite their limited coverage in existing literature and limited awareness in industry practice. Unsafe HRC can be the result of various causes. Such causes often impact both the physical and mental wellbeing of humans interacting with the HRC system. Root causes of unsafe HRC are related to how a HRC system is designed and operated, highlighting the importance of a design-led and human-centred approach to safety in HRC (see Table 2).

Similarly, the analysis of HRC safety measures demonstrated that safety precautions need be adopted at various levels in the organisation to ensure a comprehensive approach to safety. In line with the analysis of the key hazards and their root causes, cobot-related safety measures mainly address physical hazards, workspace-related safety measures address both physical and cognitive factors, and enterprise-related safety measures address mostly cognitive and organisational factors (see Table 3 and Table 4).

These results are discussed in the sections below, where hazards and safety measures are presented based on the hazard root causes.

4.1.1 **System failures and malfunctions.** Accidents with the HRC system may be caused by malfunctions in the hardware or software components. Failures of the robot, end-effector, safety system or other peripherals may expose humans to a variety of hazardous situations. Such events can be generated by the failure of mechanical or electrical components, end-effector failure (separation), failure of a safeguarding device, as well as hazards generated by multiple failures [2]. As a result, the HRC system or some of its parts may move unexpectedly leading to a loss of movement control. Having a fail-safe system structure, together with systems for collision avoidance and detection, ensures that the risks generated by failures and malfunctions are mitigated.

4.1.2 Inappropriate integration. In many cases, the level of safety is not determined by the robot or other components individually, but the overall implementation. The inappropriate integration of HRC systems can cause a variety of hazardous situations which can impact both the physical and mental wellbeing

Table 2. Human-Robot Collaboration Hazards

Root cause	Hazard group	Hazard items	References	Harm	HF/E Domain
System failures and malfunctions	Failures or malfunctions generating hazardous contact with the cobot system	Unintended motion of any axis, movement of external axes or parts, sharp tool rotation, unexpected release of stored energy, unintended end-effector activation, and failure of safeguards or end-effector components.	[2]	Crushing, shearing, cutting, impact, abrasion, pinch points	Physical
	Failures or malfunctions generating hazardous contact with other objects	Debris generated by process operations; contact between fixtures; contact between end-effector and any fixed object; contact between cobot arm and any fixed object; materials and products falling or ejection	[2], interviews	Crushing, shearing, cutting, impact, abrasion	Physical
Inappropriate integration	Inappropriate cobot selection	Robot morphology; robot characteristics/type (e.g. force, torque, acceleration, power)	[22, 23], interviews	Stress, mental strain	Cognitive
	Inappropriate end-effector selection	Characteristics of the end-effector; movements of the end-effector; safety measures embedded in the end-effector	[2, 23], interviews	Pinch points, cutting, shearing, abrasion	Physical, Cognitive
	Inappropriate design of interaction mechanisms	Poorly designed HMI screen or operator panel; poorly designed enabling devices; inappropriate location of identification controls; inappropriate location of components that require access (troubleshooting, maintenance, adjustment)	[2, 24–27], interviews	Mental strain, stress, unhealthy postures or excessive effort, fatigue	Physical, Cognitive
	Inappropriate design of the cobot workspace	Poorly designed loading/unloading post; obscured hazards, inadequate, blocked lighting; impossibility of exiting the cobot workspace; cobot system accessible by unauthorised personnel; lack of safeguarding devices/safety measures	[2, 23], interviews	Crushing, shearing, cutting, impact, abrasion	Physical, Cognitive

Root cause	Hazard group	Hazard items	References	Harm	HF/E Domain
Inappropriate task application	Inappropriate design and programming of task application	Speed, distance, warnings of motion; overriding safety measures to increase productivity; cobot trajectory; time during collaboration; task complexity	[28–30], interviews	Mental strain, stress, impact	Physical, Cognitive
	Inappropriate testing of task application	Process-related radiation and debris; lack of virtual boundaries for mobile platforms; insufficient safety measures on end-effectors; lack of consultation with specialised workers	Interviews	Crushing, shearing, cutting, impact, abrasion, pinch points	Physical
Lacking workforce awareness	Lack of trust and acceptance	Fear of losing contact with colleagues; loss of agency; fear of job loss	[24, 31–35], interviews	Mental strain, stress	Cognitive, Organisational
	Overconfidence in the cobot system	Bypassing safety measures; assuming cobots have common sense; misuse by humans	[23], interviews	Mental strain, stress; impact	Cognitive, Organisational
Unauthorised system access	System access with malicious intent	Cyber-attacks; sabotage	[28, 30, 36, 37], interviews	Crushing, shearing, cutting, impact, abrasion, pinch points	Organisational
	Data collection and privacy	Users data collected without consent and/or awareness	Interviews	Mental strain, stress	Organisational

Table 3. Human-Robot Collaboration Safety Measures

Category	***Safety measure***	***Description***
Cobot	Cobot type	Compared to industrial robots, cobots offer active and/or passive compliance and lightweight design [38]. These features combined with low moving masses are considered inherently safe for human-robot interaction.
	Cobot appearance	While the heavy, stiff, and rigid design with potentially disclosed actuator and wires of traditional industrial robots can make humans feel uncomfortable or distressing, the design of cobots emphasises lightweight and highly integrated mechatronics with fewer pinch points and smooth surfaces [38, 39].
Cobot system	Fail-safe system structure	Integrated measures may fail and thus, it is necessary to consider complementary protective measures in the system structure to reduce the risk of harming the operator while working within the collaborative workstation [40]. Any detected failure in the safety-related parts of the control system shall result in a protective stop [2, 7].
	Tool/design operation	The mechanical design of a manipulator has a huge impact on system safety [41]. Since the tools are attached to the robot's end-effector, often procured independently of the robot, the freedom to decide on the design and approach for integration into the overall system offers a decisive margin.
	Collision avoidance	The concept of contact avoidance entails ensuring the safety of operators by preventing hazardous contacts through preventive methods and systems [39, 42–44].
	Collision detection and mitigation	The concept of contact detection and mitigation deals with the reduction of collision energy in the event of an unintended or unexpected human-robot contact to ensure the operator safety [42].
	Situational awareness	Situational awareness during collaboration is essential for humans' safety. Situational awareness enables robots to autonomously perceive and interpret their environment, thereby informing humans of potential hazards to maintain safety in collaborative settings. [45].
	Intuitive cobot programming	The communication of the human's intention and the correct interpretation of the information from the robot's perspective is a crucial factor [24, 27, 39, 41, 46].
	Work cell design	Considering the design of work cells can minimise the risks of hazardous collision and debris [2]. Light as well as graphic signs and markings on the floor visually remind workers of the importance of maintaining distance from actively operating cobots.
	Human-friendly work distribution	To reduce risks to operators in human-robot collaboration, it is important to delegate work appropriately, considering the potential of both parties, with a primary attention to the physical demand and ergonomic comfort of the operator [47].
	Human-friendly workplace arrangement	The workplace should be arranged to avoid potential physical harm caused in situations in which collisions between human and robot may occur due to insufficient space to move around [2, 48]. Appropriate ergonomic design therefore should allow for the human to avoid such contacts.

Category	*Safety measure*	*Description*
	Risk assessments	A critical process that ensures workplaces have a holistic understanding of the possible risks and harms that can occur when working with cobots and provide comprehensive strategies to mitigate hazards [2, 7].
	Simulation	Simulation-based testing enables the identification of hazards [49]. This is achieved by virtually visualising and assessing the intended programmed operation.
	Physical testing	Testing at a slower speed, without end-of-arm tooling, and without a workpiece allows to assess the cobot for unexpected movements and collisions that may exist in the workplace that cannot be accounted for in a simulation. Further tests allow to evaluate the end-effector, other equipment and the workpiece.
Enterprise and context	Training to build knowledge and skills	It is important to consider the various competencies, skills, and knowledge that different stakeholders require in order to be adequately prepared to work with collaborative robots [35, 50].
	Training to improve acceptance	Training programs can also be used to promote a safe, healthy and inclusive working environment [51], which can help cobot users and operators feel more comfortable working with cobots and increase their level of trust [52].
	Assistive technology for training	An emerging opportunity to prepare the operators before they come into contact with their physical robotic counterparts is given through the use of virtual or augmented reality [27].
	Supporting worker agency	The introduction of cobots changes the structure in the socio-technical system, thus reducing workers' agency [35]. Consultation with specialist staff and technician in the development of new collaborative human-robot tasks can provide operators with capacity-building opportunities that present a path for how their skills and knowledge can grow alongside changing industries.

Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002

Table 4. Hazards vs. Safety Measures in Human-Robot Collaboration

		Failures and Malfunctions		***Inappropriate Integration***				***Inappropriate Task Application***		***Unauthorised System Access***		***Lacking Workforce Awareness***	
		Failures or malfunctions generating hazardous contact with the cobot system	Failures or malfunctions generating hazardous contact with the other objects	Inappropriate cobot selection	Inappropriate end-effector selection	Inappropriate design of interaction mechanisms	Inappropriate design of cobot workspace	Inappropriate design and programming of task application	Inappropriate testing of task application	Access with malicious intent	Data collection and privacy	Lack of trust and acceptance	Overconfidence in the cobot system
HF/E Factors		*Physical*	*Physical*	*Cognitive*	*Physical, Cognitive*	*Physical, Cognitive*	*Physical, Cognitive*	*Physical, Cognitive*	*Cognitive*	*Cognitive, Organisational*	*Cognitive, Organisational*	*Organisational*	*Organisational*
Cobot	Cobot type	x	x	x									
	Cobot appearance			x									
Cobot System	Fail-safe system structure	x	x							x			
	Tool/design operation	x	x		x								

Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002

	Collision avoidance	x	x	x									
	Collision detection and mitigation	x	x	x									
	Situational awareness					x							
	Intuitive cobot programming					x							
	Work cell design						x						
	Human-friendly work distribution							x					
	Human-friendly workplace arrangement						x						
	Risk assessments								x				
	Simulation								x				
	Physical testing								x				
Enterprise & Context	Training to build knowledge and skills	x	x									x	x
	Training to improve acceptance										x	x	x
	Assistive technology for training											x	x
	Supporting worker agency											x	x

of humans. In general, HRC can potentially improve working conditions for operators by providing several benefits including improved ergonomics - the reduction of physical and mental loading. For example, in a collaborative assembly scenario, HRC reduces the risk of strain injuries due to the lower physical effort [53]. To ensure such benefits occur, particular attention should not focus exclusively on the robot itself, but also on related parts, such as safety system, end-effectors, human-machine interface, and the workpiece. This highlights the importance of a design-led and human-centred approach to safety in HRC.

Selecting the appropriate robot ensures the physical and psychological safety of operators. The type of robot and its key characteristics, such as force, torque, acceleration, and power, can have a decisively impact on safety and human physical and mental wellbeing [23]. Compared to industrial robots, cobots offer active and/or passive compliance and lightweight design [38]. These features combined with low moving masses are considered inherently safe for human-robot interaction. While the standards and the literature clearly differentiate between industrial robots and collaborative robots, the industry has progressively blurred this line by introducing products that convert existing industrial robots for collaborative use. Despite of the availability of robots inherently capable of collaborative work, human-robot collaboration can be achieved via any kind of industrial, professional, personal service or even managerial robots [29, 54]. One integrator explained: "*I do know of some products that can be installed on industrial robots that make them behave like cobots. Where if the capacitive pads make contact with the human, they basically stop instantly, it's almost like an e-stop*". Traditional industrial robots must be equipped with adequate additional features before being capable of safe human interaction. This includes, for instance, additional software packages such as Dual Check Safety technology, the Safe Operation or SafeMove solutions of the robot manufacturers FANUC, KUKA or ABB respectively [40]. External sensors and safety equipment may also be installed so that the robot complies with one or more of the four collaborative operation modes defined by the ISO 15066. In addition to the risk for serious physical harm, an unfamiliar robot appearance can negatively affect operators' perception and contribute to a sense of insecurity and discomfort. If familiar design elements including overall form and eyes are missing, operators may feel uncomfortable and insecure when working next to them [33]. In general, it has been found that acceptance for cobots increases with higher similarity to human appearance [22].

The quintessence of any robotic system lies in its ability to interact with the environment, which is enabled by tools mounted on the end-effector. An end-effector is a device or tool that can be attached the end of a robotic arm that enables it to interact with its environment. The mechanical design of a manipulator has a huge impact on system safety [41], and the movement of end-effectors can generate hazardous situations [23], causing pinch points or injuries such as cutting or stabbing [2]. Since the tools are attached to the robot's end-effector, often procured independently of the robot, the freedom to decide on the design and approach for integration into the overall system offers a decisive margin. The interviews identified a broad range of harms that can only occur with specific end-effectors and end-of-arm tooling. A diversity of robot applications that fashion non-collaborative tools to work with end effectors, examples included hot-wires, pneumatic drills, linishers, laundry folding, and even making pancakes. The primary reason that users would customise tools and use industrial end-effectors was reported to be cost-effectiveness as purchasing collaborative specific tooling is often expensive.

Together, the robot and the end-effector closely interact with operators to enable the collaboration. Poorly designed interaction mechanisms typically enabled by devices such as human-machine interfaces (HMIs) or operator panels represent a potential hazard. Inappropriate location or identification of controls, as well as the inappropriate location of components that require access (troubleshooting, repair and adjustments) [2] may expose humans to physical hazards, as they may not be able to interact with the cobot system as intended. More broadly, complicated interaction mechanisms can have a negative impact on situation awareness, highlighting the relevance of clear interaction mechanisms. An interview participant described working with a collaborative robot as though it is *"like driving a car from outside"*, a strenuous task that can impact situational awareness. Due to factors such as unclear interaction mechanisms, the operator could have doubts and concerns about the anticipated moves of the robot [26]. Thus, the operator may not be able to identify problems and may take incorrect or unnecessary actions, which can increase the severity of harm. Clear interaction mechanisms ensure that humans can communicate with robots intuitively while understanding its intentions and movements. On the one hand, defining inputs or programming the robot should be intuitive and easy for workers. On the other hand, the information provided as feedback by the robot should be presented in an easily interpretable way to workers, so that they can have clear awareness of the system at any time [27]. Sometimes the status of a robot is presented in a form or code that is not easy to interpret for workers that do not have a high level of expertise. Typically, humans would naturally communicate by using a combination of voice and gestures, and this allows them to convey information that can be either complementary or redundant. For example, an operator could say *"take this!"* while pointing at a specific object [24]. In many cases, human actions need to be communicated by pressing buttons, which are not always within close proximity [25].

More generally, the design of the overall collaborative workspace is as important as the HRC system in terms of safety. This involves the integration of various equipment as well as broader design choices. Appropriate integration of surrounding equipment represents a significant aspect of HRC. As a collaborative workstation may consist besides the robot itself of several external devices such as the sensors, tooling, additional machines or monitoring equipment, their fusion to a functional system is highly important to guarantee the operator's safety. However, such integrated measures may fail and thus, it is necessary to consider complementary protective measures in the system structure to reduce the risk of harming the operator while working within the collaborative workstation [40]. The greatest risk is associated with the integration of general non-safety-rated devices.

In addition to the equipment selection, poor design choices for operation and maintenance can create ergonomics hazards [23]. The location of specific elements of the HRC system should allow easy accessibility, not only for regular operation but also for troubleshooting, repair, and adjustments [2]. For example, during the interviews it was noted that malfunctioning safety measures such as damaged sensors were difficult to identify and assess. While collaborative robots have their own internal diagnostics, one user claimed that they were unable to perform manual checks to assess if the safety sensors were working "*because the sensors are embedded within the machine itself*". Other examples of inappropriate location for elements of the HRC system include obscured hazards or physical obstacles in front of safety devices such as sensors or cameras impeding their correct functioning [23], inaccessible location or identification of controls (e.g. stop buttons, control panel) [2], and unsafe operator's location (e.g. working under a heavy payload cobot) [23].

Finally, in the design choices related to the working area, safety distance limit in the robot trajectory should not be tight [23], and workspace should be clearly identified. In fact, potentially hazardous movements of the HRC system may create debris resulting in an object being thrown across the workspace. During interviews, light curtains were consistently reported as a safety peripheral device that prevented workers from accidental collisions. Furthermore, graphic signs and markings on the floor visually reminded workers of the importance of maintaining distance from actively operating cobots.

4.1.3 Inappropriate task application. An appropriate integration of the HRC system is often not enough to guarantee safety. A safe design of the task application is also a key element. The operational characteristics of the task application may represent a danger in itself or may create other hazards elsewhere. Programming and testing the task application can impact operators' wellbeing significantly.

When operating collaborating robots, humans must feel comfortable, and the mental strain associated with tasks has to be bearable. Unpredictable motions of the robot can cause unpleasant reactions such as fear, shock, or surprise, and the anticipation of the potential for unexpected movements and collision may make users nervous [29]. In general, to create safe collaborative environments, cobots are intentionally limited to slower movement speeds and lower payloads to minimise

Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002

the severity of hazardous collisions. Speed, distance and warnings of motions directly influence the psychological state of operators [30]. When these safety measures are overridden to increase productivity, it directly impacts an operator's sense of comfort when using the machine. The interviews highlighted how speed and the payload assigned to collaborative robots made even well-versed users nervous to work collaboratively with them. When collaborative robots reach their maximum payload, end users and integrators noted that performance was inconsistent and unreliable. One integrator observed that once collaborative robots reach their maximum payload they tend to *"jitter a little bit, as if it's moving on its own and it gets really slow"*. When collaborative robots are a part of a larger, more complex system, these unreliable speeds and unexpected movements can impact entire work processes.

Programming alone is not sufficient to guarantee safety. Comprehensive testing together with the consultation with specialised workers are essential. Integrators can overlook the potential risks associated with task-specific risks when they do not consult workers with specialist knowledge (e.g. welders) during the implementation process of the task. Appropriate testing allows to highlight hidden risks, which may have been overlooked during the initial integration of the HRC system. Such hidden risks can be related to specific components of the system, such as end-effectors of the platform on which they are installed, as well as specific aspects of the process.

As mentioned earlier, end-effectors often represent a source of hazards. The interviews revealed that many tasks have been developed ad-hoc by users without risk assessments or consultation with integrators. A strong example was highlighted in the interview study where a user attached a drill to the end effector using plastic zip ties. Upon reflection the research participant realised that this was hazardous as they would have been unable to immediately turn off the drill if something had gone wrong. In general, end effectors were often reported as missing sufficient safety measures such as sensors installed onto them. This is especially important considering that safety sensors are one of the main ways that physical harm and collisions are prevented. Furthermore, research participants highlighted that when testing applications using simulation software, end-effectors were simulated as static objects. This limits users from identifying errors or issues that may cause physical risks before a full operational run.

In addition to end-effectors, another essential component for which risks can be overlooked is the platform on which robots are installed. A key selling point for collaborative robots is that they are lighter and smaller compared to industrial robots, making them easier to move and transport. The interviews identified a multitude of ways end users installed collaborative robots upon mobile equipment such as trolleys, Automated Guided Vehicles (AGVs), or Autonomous Mobile Robots (AMRs). Once a cobot is installed on mobile equipment, the work cell becomes dynamic and not clearly defined. For end users who transported a cobot to

Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002

a variety of settings, the greatest potential for collision was in the initial set-ups, when safety measures and sensors are either turned off or require reconfiguration and recalibration for the new setting. This stresses the importance of considering the dimension of workspace and/or cell design in minimise the potential risk of hazardous collisions.

An attention to the application as a whole also allows to reduce process-related risks such as the generation of dangerous debris created by the task at hand (e.g. metalwork, linishing), which may expose humans to process-related radiation [2], and the potential risk of pinch points. Debris can cause unsafe working conditions and injuries for humans including but not limited to; poor ventilation, tripping & slipping hazards, fires, chemical burns, and spills. For these processes, debris is a natural by-product of the task. This requires organisations to consider non-cobot related risk mitigation strategies to safely manage debris. Debris and dust created by task applications may also accumulate in the open joints of a cobot arm, causing component deterioration. Another common form of debris mentioned by a supplier was the risk of cobots accidentally spilling the contents of a container or a box breaking while a cobot is carrying it. Similarly, pinch points can be the result of unsafe process design, and they can often be detected with accurate testing. An integrator noted that historically, testing kits would include fake fingers that could be used to measure severity and minimise the risk of pinch points. However, they remarked that this had recently become less common and highlighted that fake fingers were not a standard requirement for risk assessments.

4.1.4 Workforce awareness. The lack of workforce awareness on various levels can create hazards to the physical and mental wellbeing of humans interacting with the HRC system.

The lack of understanding about the way in which collaborative robots operate can generate a variety of undesired consequences.

First of all, it is important to consider the various competencies, skills, and knowledge that different stakeholders require in order to be adequately prepared to work with collaborative robots [35]. A lack of knowledge and experience in operating cobots was largely attributed as the leading cause of increased physical and psychological risks, according to most research participants. Understanding this, it is clear that training and short courses are critical to ensuring that users who are programming, operating, or maintaining cobots remain safe [29]. Such lack of knowledge may also generate overconfidence in the system within operators. Often, due to the collaborative nature, operators assume collaborative robots has ‘common sense’ expecting a robot to move in a predictable human-like manner. When a robot needs to move from A to B, regardless of whether the two points are physically close to each other, it can sometimes take roundabout paths to reach the destination. This issue was a primary concern for an end user who trains students to use collaborative robot; they remarked that they remind students that *"it's still a robot and they need to move out of the way, especially when they're running things for the first time"*.

On an organisational level, the lack of understanding about how the introduction of HRC affects the overall work structure may generate stress, lack of trust, and fear of job loss. As the introducing HRC into the workplace affects the social environment of a workplace, some workers may fear that they might lose contact with their colleagues [31, 32]. The introduction of HRC also triggers changes in the workers' roles. While most interview participants appeared excited about the future of upskilling workers to operate collaborative robots, there seemed little consideration to the loss of artisanal knowledge and skills that may occur. HRC can also have impact upon an operator's agency in feeling a sense of ownership and responsibility over their work [35]. This may contribute to a devaluing of the knowledge and skills that they may possess. In interviews with a supplier, they highlighted that in manufacturing settings programming task applications for collaborative was typically conducted by more senior personnel. Operators were reported to simply turn collaborative robots on and off and the beginning and end of their shifts. It was reported in the interviews that operators enjoyed the ability to easily tailor programming to mould to their working style and knowledge of specialist task applications. This indicates that a lack of operator engagement in programming may contribute to a lack of worker agency and overall acceptance of HRC in the workplace. More generally, the introduction of HRC can induce fear of job losses among workers [24, 33]. Accordingly, during interviews, several integrators and suppliers commented that when HRC was introduced to a new site most workers were cautious to engage.

4.1.5 Unauthorised system access. In the context of advanced manufacturing systems, where devices and machines are interconnected, safety often relates to security, as the system could be vulnerable to events such as cyber-attacks that induce unwanted behaviours [30, 36]. Malicious access may be represented by not only external cyber-attacks, but also sabotaging actions. In the context of cyber security, collaborative cyber-physical systems include a variety of features, including hardware, sensor network, and information and communication technologies. This allows to connect these systems to their intra- or internet, thus exposing them to security risks [37]. As a result, security and safety aspects become strictly related. In some cases, the potential impact that a cyber-attack can have is wider as it happens on a systemic level. Cyber-attacks can affect many robotic systems or entire manufacturing sites at once, while in most cases, hazardous events such as collisions affect a single operator.

Malicious access can also occur internally from operators. During interviews, an integrator shared various examples of how operators in the past had changed programming to encourage collision. He noted that in one case "*they've purposely* *gone in and sort of tried to change the robot code so it crashes into something on purpose, jammed up a conveyor or destroy sensors installed*". Another integrator stated that it was a relatively common occurrence to hear of operators sabotaging HRC systems as they saw it as a threat to their livelihood. The integrator explained that sabotaging HRC systems could be a relatively easy task that would not require any programming or technical knowledge, as "*you can just go to the teach pendant and just delete a few lines here and there, and you wouldn't really need to know what those lines even meant*". Intentionally damaging sensors was also disclosed as a sabotaging method. One research participant disclosed that they had heard of operators cleaning robots with caustic solutions and abrasively handling the robots.

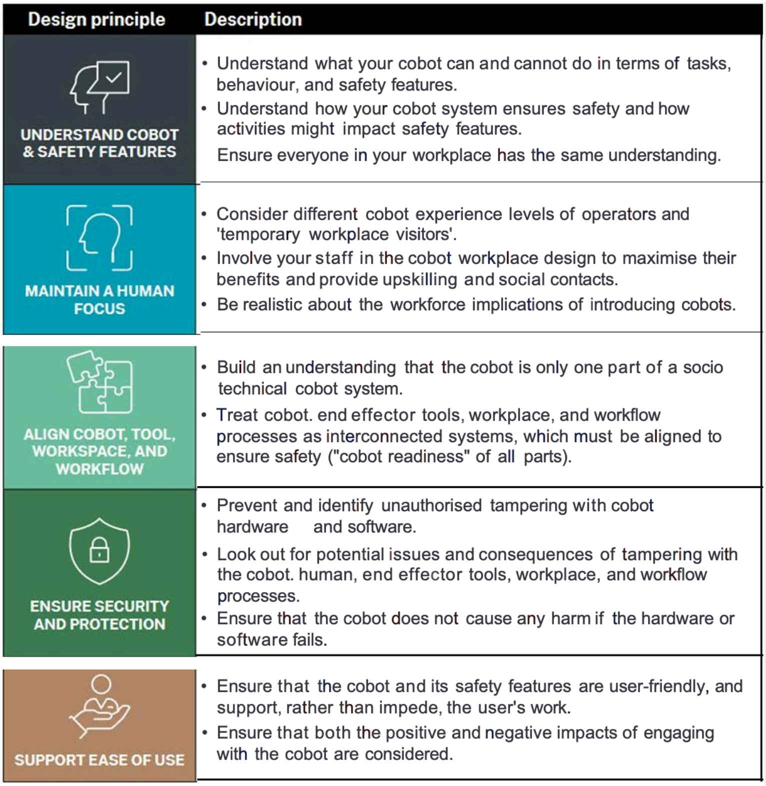
Other forms of data access are represented by the collection of data about the HRC system, including the operator, without their consent or awareness. Considering that robots can capture an array of data from their safety systems, there is a risk that operators and user data may be collected, used, and sold without user consent. In the interview study it became clear that many organisations in the industry were already interested in the potential value of this data in the development of future products and services.

4.2 Design principles for safe Human-Robot Collaboration

The literature review and interview study confirmed the challenge of providing universally applicable safety measures due to the large variety of HRC and associated workplaces. Some measures might work well for specific HRC cases, while others need larger adjustments, and others might not be suitable at all. In addition, interview participants from different backgrounds stated that existing standards tend to be inaccessible to many cobot users. Especially their highly technical wording hinders users to understand and address risks appropriately, preventing them from complying with the standards. Another issue is also the strong focus of physical hazards and harms, with limited attention to cognitive and organisational factors.

The identified design principles provide overarching guidance in ensuring safe HRC from a holistic socio-technical perspective. Their abstract nature enables their applicability for various HRC, types of human-cobot workplaces and across the entire life-cycle of a cobot and workplace. They can form the basis of organisation-, workplace- and HRC-specific safety measures and guidelines. Fig. 5 shows the five design principles of Understand cobot and safety features, Maintain a human focus, Align cobot, tool, workspace and workflow, Ensure security and protection, and Support ease of use, which are explained in the following.

Interaction Design and Architecture(s) Journal - IxD&A, N.61, 2024, pp. 66 - 97
DOI: 10.55612/s-5002-061-002



Design principle	Description
UNDERSTAND COBOT & SAFETY FEATURES	• Understand what your cobot can and cannot do in terms of tasks, behaviour, and safety features. • Understand how your cobot system ensures safety and how activities might impact safety features. Ensure everyone in your workplace has the same understanding.
MAINTAIN A HUMAN FOCUS	• Consider different cobot experience levels of operators and 'temporary workplace visitors'. • Involve your staff in the cobot workplace design to maximise their benefits and provide upskilling and social contacts. • Be realistic about the workforce implications of introducing cobots.
ALIGN COBOT, TOOL, WORKSPACE, AND WORKFLOW	• Build an understanding that the cobot is only one part of a socio technical cobot system. • Treat cobot. end effector tools, workplace, and workflow processes as interconnected systems, which must be aligned to ensure safety ("cobot readiness" of all parts).
ENSURE SECURITY AND PROTECTION	• Prevent and identify unauthorised tampering with cobot hardware and software. • Look out for potential issues and consequences of tampering with the cobot. human, end effector tools, workplace, and workflow processes. • Ensure that the cobot does not cause any harm if the hardware or software fails.
SUPPORT EASE OF USE	• Ensure that the cobot and its safety features are user-friendly, and support, rather than impede, the user's work. • Ensure that both the positive and negative impacts of engaging with the cobot are considered.

Fig. 5. Design principles for safe human-robot collaboration

4.2.1 Understanding cobot and safety features. *Understanding cobot and safety features* includes an understanding of what a cobot can and cannot do in terms of tasks, behaviour and safety features as well as an understanding of how your cobot system ensures safety and how activities might trigger unwanted safety features.

Collaboration is a vague term that can be interpreted in various ways. In the interview study most cobot users did not use their cobot for collaborative applications. Instead, most cobots were used for co-existent and cooperative functions with human workers. Research participants that were responsible for selling cobots often stated that the main selling point is that cobots are cheaper, easier to use, and take up less space on factory floors compared to industrial robots. It appeared that several cobot user companies were interested in the technology for creating fenceless applications that free up expensive floor space, rather than for collaborative applications. Accordingly, a safety peripheral manufacturer that converts industrial robots for collaborative use explained that for many of their clients, the desire to purchase their equipment was so that they could go fenceless. Fences they explained were “a hindrance to good flow through”.

Key to a safe collaboration is an understanding of the other side. In the first place, this addresses the cobot, its characteristics, and features. A lack of understanding of what the cobot can do and how it will behave can limit both cobot performance and safety. For instance, a lack of knowledge of the maximum payload could result in handling a too heavy object, which might result in jittery or no movements. In addition, cobot speed needs to be considered and adjusted when defining a collaborative application, along with considering the possible position accuracy to decide which applications are feasible. Another example is the knowledge about possible singularities of cobot manipulators when all joints and segments form a straight line, which can confuse unexperienced operators [55].

Similarly, understanding a cobot's safety features is crucial. On the one hand, this means knowing the safety capabilities and features to avoid wrong expectations. This could include the fact that even though a cobot itself might be inherently safe, adding a wrong end-effector, such as a knife or welding torch, could jeopardize those. On the other hand, it also means roughly knowing how the safety features work: for instance, collision impact could be reduced by more flexible joints, which would lead to reduced position accuracy; or proximity sensors could be triggered by visitors or reflective surfaces, which would cause frequent cobot stops and frustration of the operator's side.

This understanding should not be limited to operators alone but, with varying degrees, also needs to cover others, such as potential workplace visitors or staff passing through a cobot workplace, production managers to efficiently plan and manage production processes and performance, and senior managers responsible for investment decisions.

4.2.2 Ensuring a human focus. *Ensuring a human focus* includes considering different cobot experience levels of operators and 'temporary workplace visitors' as well as involving staff in the cobot workplace design to maximise the benefits for them and provide upskilling and social contacts. The human focus means paying attention to all human factors, including physical, cognitive, and organisational ergonomics. Our study has highlighted a general lack of focus on psychological and social hazards generated by the introduction of HRC. In order to minimise such risks, it is important to include broader safety measures at the enterprise level. The lack of workforce awareness about the benefits and implications of HRC can be mitigated by various forms of training. Training to build knowledge and skills ensures that users who are programming, operating, or maintaining cobots remain safe [29]. In addition to this, training should be use to promote acceptance and allow workers' agency. A lack of knowledge and experience in operating cobots was largely attributed as the leading cause of increased physical and psychological risks, according to most research participants. Alongside the physical design of cobot, it is important to consider training to helps cobot users and operators feel more comfortable working with cobots. Comfort is based upon predictability and familiarity, both of which minimise the mental strain that operators experience when they are fearful of cobots. Managers play a critical role in supporting their staff by actively working to maintain the agency of their workers. One way that worker agency can be supported is by encouraging staff to optimise their work assignments to their working preferences and to explore how else a cobot can be used. A broader approach to supporting worker agency is consultation and co-designed solutions with operators. Consultation with specialist staff and technician in the development of new collaborative human-robot tasks can provide operators with capacity-building opportunities that present a path for how their skills and knowledge can grow alongside changing industries.

4.2.3 Aligning cobot, workspace and workflow. *Aligning cobot, workspace and workflow* includes building an understanding that the cobot is only one part of a socio-technical cobot system and, as a result, treating cobot, end-effector tools, workplace and workflow processes as an interconnected system, which needs to be aligned to ensure safety ("cobot readiness" of all parts). The root cause analysis conducted for this study has highlighted the importance of safe integration and safe design of task applications in HRC. Overall, this demonstrates the need for a systemic approach in the design of collaborative systems. A design-led approach means that safety should address all levels of HRC – cobots, cobot system, and enterprise and context – and how all the elements interact within and across levels (see Fig. 2).

4.2.4 Ensuring security and protection. *Ensuring security and protection* includes preventing and identifying unallowed tampering with cobot hardware and software as well as looking out for potential issues and consequences of tampering on the cobot, human, end-effector tools, workplace and workflow processes. Malicious access could occur from either internal or external actors. At the enterprise level, human operators or other staff could attempt to actively sabotage the cobot system. At a broader level, connecting cyber-physical systems comprised of a variety of features, including hardware, sensor network, and information and communication technologies may expose them to security risks [37]. Regardless of where the unauthorised originated from, attacks may cause the system to behave unexpectedly. A design-led approach should consider these eventualities and ensure that the system is secure and safe.

4.2.5 Supporting ease of use. *Supporting ease of use* includes ensuring the cobot and its safety features are user friendly and support the staff's work, and that the positive and negative impact of engaging with the cobot is considered. When programming a cobot, the communication of the human's intention and the correct interpretation of the information from the robot's perspective is a crucial factor [39, 41]. It has been found that in practical industrial applications, the programming of the robot consumes a large portion of the human worker's cognitive interaction, as humans have to provide the robot with explicit motion-oriented instructions [27]. While traditional programming techniques such as lead-through or coding tend to be quite unnatural, new user interface strategies are emerging that are more closely aligned with a person's native communication channels.

5 Conclusions

The characteristics of collaborative robots bring together a unique combination of social and technical dimensions, calling adopting a design-led approach to HRC. The key contribution of this study lies in the development of principles for safe HRC. These have been formulated based on a socio-technical analysis of hazards, taking into account physical, cognitive, and organisational factors, and by comparing them with existing safety measures to identify potential gaps. By collecting data from both high-quality scholarly sources and the industry, this study provides contribution to academia and industry practice.

The analysis of hazards and harms has concluded that while that existing standards and practices focus predominantly on mitigating physical risks, HRC is a complex socio-technical system that can potentially harm the physical and mental wellbeing of humans. In line with these results, most existing safety measures are focussed on preventing and mitigating physical risks, with safety measures at the enterprise and context level being almost exclusively addressed in the interview study (e.g. supporting worker agency, training to improve acceptance). Based on the results provided by the background research, the design principles for safe HRC aim to provide overarching guidance on what entails and supports a safe collaborative workplace. The background research, and the interview study in particular, highlighted the confusion around what a cobot is, how it can be used and how it can have a critical impact on safety. Therefore, the first design principle '*Understanding cobot and safety features*' focusses on understanding what a collaborative robot can do in terms of tasks, behaviour, and safety limits. The background research also highlighted the predominant focus on the physical harm that robots and cobots may inflict. Thus, the second design principles '*Ensuring a human focus*' ensures that safety is considered as a comprehensive term that includes not only the physical aspects that standards and existing guidelines traditionally focus on, but also the broader cognitive and organisational considerations that impact workers. Maintaining the same holistic perspective, the background research also found that hazards may be generated by a variety of components and applications of the cobot system. Thus, the third design principle '*Aligning cobot, workspace and workflow*' emphasises the socio-technical complexity of HRC, noting that cobot, end-effector tools, workplace, and workflow processes should be managed as an interconnected system. Based on the challenges emerging from the increased inter-connectivity of workplaces across various industries, the fourth design principles '*Ensuring security and protection*' emphasises the need for identifying and preventing any type of unauthorised access, from sabotage to cyber-attacks. Finally, the fifth design principle '*Supporting ease of use*' highlights the need for HRC to support humans' work and ensuring that their use is intuitive.

The abstract nature of the design principles ensures their applicability across various workplaces and lifecycles. This contribution is substantiated by some of the findings of the interview study, where research participants noted that standards are not always easily accessible due to the highly technical language. In this regard, the intent of the design principles is not to replace existing standards, but to provide support with a practical and design-led approach to safety in HRC.

Overall, the design principles for safe HRC can inform future research about new guidelines for safe human-robot collaborative work.

Acknowledgments. This research was approved by the University of Technology Sydney Ethics Committee (UTS HREC REF NO. ETH21-6244) and is part of The Impact of Robotics on Work Health and Safety research project funded by the Centre for Work Health and Safety, NSW Department of Customer Service. The authors would like to thank the project team and the participants for sharing their experience and their valuable input.

CRediT author statement. Laura Tomidei: Methodology, Formal analysis, Writing – original draft preparation. **Matthias Guertler:** Conceptualisation, Methodology, Formal analysis, Validation, Writing – Review and Editing. **Nathalie Sick:** Methodology, Validation, Formal Analysis, Writing – Review and Editing. **Gavin Paul:** Validation, Writing – Review and Editing. **Marc Carmichael:** Validation, Writing – Review and Editing.

References

1. Tchane Djogdom G.V., Otis M.J.D., Meziane R.: Dynamic time warping–based feature selection method for foot gesture cobot operation mode selection, International Journal of Advanced Manufacturing Technology, 126, pp. 4521–4541 (2023), doi: 10.1007/s00170-023-11280-w
2. ISO 10218-2: Robots and robotic devices — Safety requirements for industrial robots, Part 2: Robot systems and integration (ISO 10218-2:2011) BSI Standards Publication (2011)
3. ISO 10218-1: Robots and robotic devices — Safety requirements for industrial robots, Part 1: Robots (ISO 10218-1:2011) BSI Standards Publication, pp. Annex A-List of significant hazards (2014)
4. Haddadin S., Croft E.: Physical Human–Robot Interaction, Springer Handbooks, pp. 1835–1874 (2016), doi: 10.1007/978-3-319-32552-1_69/FIGURES/30
5. Centre for Work Health and Safety NSW, University of Technology Sydney, Centre for Inclusive Design, Kairos Now: Phase 2 report: Work health and safety risks and harms of cobots, (2021)
6. Dobra Z., Dhir K.S.: Technology jump in the industry: human–robot cooperation in production, Industrial Robot, 47(5), pp 757-775 (2020), doi: 10.1108/IR-02-2020-0039
7. ISO/TS 15066: PD ISO / TS 15066 : 2016 "Robots and robotic devices — Collaborative robots" International Standards Organisation (ISO) (2016)
8. Baxter G., Sommerville I.: Socio-technical systems: From design methods to systems engineering, Interacting with Computers, 23(1), pp. 4–17 (2011), doi: 10.1016/j.intcom.2010.07.003
9. Davies R., Coole T., Smith A.: Review of Socio-technical Considerations to Ensure Successful Implementation of Industry 4.0, Procedia Manufacturing, 11, pp. 1288–1295 (2017), doi: 10.1016/j.promfg.2017.07.256
10. Kopp T., Baumgartner M., Kinkel S.: Success factors for introducing industrial human-robot interaction in practice: an empirically driven framework, International Journal of Advanced Manufacturing Technology, (2020), doi: 10.1007/s00170-020-06398-0
11. Guertler M., Carmichael M., Paul G., Sick N., Tomidei L., Hernandez Moreno V., Wambsganss A., Amin M., Grace R., Cockburn J., Frijat L., Hussain S.: Guidelines for the Safe Collaborative Robot Design and Implementation, NSW Government: Centre for Work Health and Safety (2022)

12. Zacharaki A., Kostavelis I., Gasteratos A., Dokas I.: Safety bounds in human robot interaction : A survey, Safety Science, 127, pp. 104667 (2020), DOI: 10.1016/j.ssci.2020.104667
13. ISO 26800: Ergonomics — General approach, principles and concepts, (2011)
14. International Ergonomics Association: What Is Ergonomics?, https://iea.cc/what-is-ergonomics/
15. Stramler J.H.: The Dictionary for Human Factors, Routledge, (1993), doi: 10.1201/9780203736890
16. Tomidei L., Sick N., Guertler M., Frijat L., Carmichael M., Paul G., Wambsganss A., Moreno V.H., Hussain S.: BEYOND TECHNOLOGY-THE COGNITIVE AND ORGANISATIONAL IMPACTS OF COBOTS, in Australasian Conference on Robotics and Automation ACRA (2022)
17. Levy Y., Ellis T.J.: A systems approach to conduct an effective literature review in support of information systems research, Informing Science, 9, pp. 181–211 (2006), doi: 10.28945/479
18. DiCicco-Bloom B., Crabtree B.F.: The qualitative research interview, Medical Education, 40(4), pp. 314–321 (2006), doi: 10.1111/j.1365-2929.2006.02418.x
19. Martin-Martin A., Orduna-Malea E., Thelwall M., Lopez-Cozar E.D.: Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories, Journal of Informatics, 12, pp. 1160–1177 (2018), doi: 10.1016/J.JOI.2018.09.002
20. Karen H., Sandra J.: Contextual Inquiry: A Participatory Technique for System Design, Participatory Design. pp. 177–210. CRC Press (2017), doi: 10.1201/9780203744338-9
21. Corbin J.M., Strauss A.: Grounded theory research: Procedures, canons, and evaluative criteria, Qualitative Sociology, 13(1), pp. 3–21 (1990), doi: 10.1007/BF00988593
22. Bauer A., Wollherr D., Buss M.: Human–Robot Collaboration: A Survey International Journal of Humanoid Robotics, 5(1), pp. 47–66 (2008), doi: 10.1142/S0219843608001303
23. Khalid A., Kirisci P.T., Ghrairi Z., Pannek J.: Implementing safety and security concepts for human-robot collaboration in the context of Industry 4.0, in 39th MATADOR Conference, pp. 2–8 (2017)
24. Maurtua I., Ibarguren A., Kildal J., Susperregi L., Sierra B.: Human-robot collaboration in industrial applications: Safety, interaction and trust, International Journal of Advanced Robotic Systems, 14(4), pp. 1–10 (2017), doi: 10.1177/1729881417716010
25. Michalos G., Makris S., Spiliotopoulos J., Misios I., Tsarouchi P., Chryssolouris G.: ROBO-PARTNER: Seamless human-robot cooperation for intelligent, flexible and safe operations in the assembly factories of the future, Procedia CIRP, 23, pp. 71–76 (2014), doi: 10.1016/j.procir.2014.10.079
26. Michalos G., Kousi N., Karagiannis P., Gkournelos C., Dimoulas K., Koukas S., Mparis K., Papavasileiou A., Makris S.: Seamless human robot collaborative assembly – An automotive case study, Mechatronics, 55, pp. 194–211 (2018), doi: 10.1016/j.mechatronics.2018.08.006
27. Villani V., Pini F., Leali F., Secchi C.: Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications, Mechatronics, 55, pp. 248–266 (2018), doi: 10.1016/j.mechatronics.2018.02.009
28. Galin, R., Meshcheryakov, R.: Review on Human–Robot Interaction During Collaboration in a Shared Workspace, in Lecture Notes in Computer Science (including

subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2019), doi: 10.1007/978-3-030-26118-4_7/FIGURES/2

29. Murashov V., Hearl F., Howard J.: Working safely with robot workers: Recommendations for the new workplace, Journal of Occupation and Environment Hygene, 13(3), pp. D61–D71 (2016), doi: 10.1080/15459624.2015.1116700
30. Robla-Gomez S., Becerra V.M., Llata J.R., Gonzalez-Sarabia E., Torre-Ferrero C., Perez-Oria J.: Working Together: A Review on Safe Human-Robot Collaboration in Industrial Environments, IEEE Access, 5, pp. 26754–26773 (2017), doi: 10.1109/ACCESS.2017.2773127
31. Bröhl C., Nelles J., Brandl C., Mertens A., Schlick C.M.: TAM reloaded: A technology acceptance model for human-robot cooperation in production systems, Communications in Computer and Information Science, 617, pp. 97–103 (2016), doi: 10.1007/978-3-319-40548-3_16
32. Gervasi R., Mastrogiacomo L., Franceschini F.: A conceptual framework to evaluate human-robot collaboration, International Journal of Advanced Manufacturing Technology, 108, pp. 841–865 (2020), doi: 10.1007/s00170-020-05363-1
33. Sauppé A., Mutlu B.: The social impact of a robot co-worker in industrial settings, in Conference on Human Factors in Computing Systems - Proceedings, 2015-April, pp. 3613–3622 (2015), doi: 10.1145/2702123.2702181
34. Lewis M., Sycara K., Walker P.: The Role of Trust in Human-Robot Interaction, Studies in Systems, Decisions, Control, 117, pp. 135-159 (2018), doi: 10.1007/978-3-319-64816-3_8
35. Weiss A., Wortmeier A.K., Kubicek B.: Cobots in Industry 4.0: A Roadmap for Future Practice Studies on Human-Robot Collaboration, IEEE Transactions on Human Machine Systems, 51, pp. 335–345 (2021), doi: 10.1109/THMS.2021.3092684
36. Bi Z.M., Luo M., Miao Z., Zhang B., Zhang W.J., Wang L.: Safety assurance mechanisms of collaborative robotic systems in manufacturing, Robotics and Computer Integrated Manufacturing, 67, (2021), doi: 10.1016/j.rcim.2020.102022
37. Khalid A., Kirisci P., Khan Z.H., Ghrairi Z., Thoben K.D., Pannek J.: Security framework for industrial collaborative robotic cyber-physical systems, Computers in Industry, 97, pp. 132–145 (2018), doi: 10.1016/j.compind.2018.02.009
38. Siciliano B., Khatib O.: Springer Handbook of Robotics (2016), doi: 10.1007/978-3-319-32552-1
39. Vysocky A., Novak P.: Human - Robot collaboration in industry, MM Science Journal, 2016-June, pp. 903–906 (2016), doi: 10.17973/MMSJ.2016_06_201611
40. Magrini E., Ferraguti F., Ronga A.J., Pini F., De Luca A., Leali F.: Human-robot coexistence and interaction in open industrial cells, Robotics and Computer-Integrated Manufacturing, 61, pp. 101846 (2020), doi: 10.1016/j.rcim.2019.101846
41. Aldini S., Carmichael M.G., Liu D.: A risk reduction framework for design of physical human-robot collaboration Australasian Conference on Robotics and Automation, ACRA, 2019-Decem, (2019), http://hdl.handle.net/10453/137199
42. Gualtieri L., Rauch E., Vidoni R.: Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review Robotics and Computer-Integrated Manufacturing, 67, pp. 101998 (2021), doi: 10.1016/j.rcim.2020.101998
43. Elkmann N., Behrens R.: New approaches to improve the design of HRC robot applications (computer-aided safety), in IROS 2019 (2019)

44. Stengel D., Ostermann B., Ding H., Bortot D., Schiller F., Stursberg O., Bengler K., Huelke M., Som F., Strunz U.: An Approach for Safe and Efficient Human-Robot Collaboration, in 6th International Conference Safety of Industrial Automated Systems SIAS 2010 (2010), (2010)
45. Bavle H., Luis Sanchez-Lopez J., Cimarelli C., Tourani A., Voos H.: From SLAM to Situational Awareness: Challenges and Survey, Sensors, 10 (2023), doi: 10.3390/s23104849
46. Sutjipto S., Carmichael M.G., Paul G.: Comparison of strength profile representations using musculoskeletal models and their applications in robotics, Frontiers in Robotics and AI, 10, pp. 1265635 (2023), doi: 10.3389/FROBT.2023.1265635/BIBTEX
47. Pearce M., Mutlu B., Shah J., Radwin R.: Optimizing Makespan and Ergonomics in Integrating Collaborative Robots into Manufacturing Processes, IEEE Transactions on Automation Science and Engineering, 15, pp. 1772–1784 (2018), doi: 10.1109/TASE.2018.2789820
48. Michalos G., Makris S., Tsarouchi P., Guasch T., Kontovrakis D., Chryssolouris G.: Design Considerations for Safe Human-robot Collaborative Workplaces, Procedia CIRP, 37, pp. 248–253 (2015), doi: 10.1016/j.procir.2015.08.014
49. Huck T.P., Münch N., Hornung L., Ledermann C., Wurll C.: Risk assessment tools for industrial human-robot collaboration: Novel approaches and practical needs, Safety Science, 141, pp. 105288 (2021), doi: 10.1016/j.ssci.2021.105288
50. Kildal J., Tellaeche A., Fernández I., Maurtua I.: Potential users' key concerns and expectations for the adoption of cobots, Procedia CIRP, 72, pp. 21-26 (2018), doi: 10.1016/j.procir.2018.03.104
51. Simões A.C., Pinto A., Santos J., Pinheiro S., Romero D.: Designing human-robot collaboration (HRC) workspaces in industrial settings: A systematic literature review, Journal of Manufacturing Systems, (2022), doi: https://doi.org/10.1016/j.jmsy.2021.11.007
52. Faccio M., Granata I., Menini A., Milanese M., Rossato C., Bottin M., Minto R., Pluchino P., Gamberini L., Boschetti G., Rosati G.: Human factors in cobot era: a review of modern production systems features, Journal of Intelligent Manufacturing, (2022), doi: 10.1007/s10845-022-01953-w
53. Cherubini A., Passama R., Crosnier A., Lasnier A., Fraisse P.: Collaborative manufacturing with physical human-robot interaction, Robotics and Computer-Integrated Manufacturing, 40, pp. 1–13 (2016), doi: 10.1016/j.rcim.2015.12.007
54. Probst L., Frideres L., Pedersen B., Caputi C.: Service Innovation for Smart Industry EU Business Innovation Observatory (2015)
55. Carmichael M.G., Liu D., Waldron K.J.: A framework for singularity-robust manipulator control during physical human-robot interaction, The International Journal of Robotics Research, 36(5-7), pp. 861–876 (2017), doi: 10.1177/0278364917698748