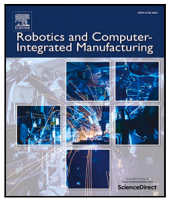
Robotics and Computer–Integrated Manufacturing 78 (2022) 102386

Contents lists available at ScienceDirect

Robotics and Computer-Integrated Manufacturing

journal homepage: www.elsevier.com/locate/rcim



Full length article

Deliberative safety for industrial intelligent human–robot collaboration: Regulatory challenges and solutions for taking the next step towards industry 4.0

Atieh Hanna [a,*], Simon Larsson [b], Per-Lage Götvall [a], Kristofer Bengtsson [a]

[a] Volvo AB, Group Trucks Operations division, Quality and Engineering function, Manufacturing Technologies department, Research and technology development group, Gothenburg, Sweden

[b] University of Gothenburg, School of Business, Economics and Law, Gothenburg Research Institute, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Collaborative robots
Automation
Industry 4.0
Safety
Safety Standards
Regulation
Machinery Directive

ABSTRACT

In previous scholarly literature, safety is understood as a main obstacle for introducing human–robot collaboration in industrial production. This interdisciplinary paper is concerned with the safety and regulation of human–robot collaboration and contribute to this debate through a case study of stakeholders in Sweden, exploring the views of the involved stakeholders which is largely absent in previous research literature. The case study concludes that while stressing some potential benefits, stakeholders within the industry are generally reluctant to human–robot collaboration. Current regulation and safety standards are understood to be one of the prominent obstacles to such solutions. Based on the perspectives of the stakeholders as well as an analysis of current regulation and safety standards, the paper identifies the following problems with current regulation: (i) existing categories and conceptualizations used to guide safety evaluation are problematic, (ii) intelligence and autonomous aspects of collaborative systems are not sufficiently addressed, (iii) current standards do not enable evaluation of the trade off between safety, efficiency and flexibility, and (iv) the regulation has a lack of focus on active safety and using the control system as a safety measure.

In an attempt to address these identified problems, the difference between traditional collaborative robots and intelligent human–robot collaboration is analyzed in the paper and a new safety approach is suggested, called *Deliberative safety*, which allows the humans and robots to switch between different safety measures based on the need for flexibility or efficiency to reach production goals. While considering system performance, we propose a taxonomy to better support the design of deliberative safety as well as five safety measures to use in a deliberative safety approach. These measures include available measures like perimeter safety, zone safety and reactive safety to more advanced measures like planned and active safety, and when used together, they can enable intelligent human–robot collaboration.

1. Introduction

Collaborative robots, so-called *cobots*, are described as the *frontier of industrial robot* [1], and it is argued that these robots can create competitive industries by increasing flexibility, efficiency, and quality in industrial production [2]. Both the producers of cobots and scholars (e.g. [3]) emphasize that cobots can make production more flexible or agile responding to shorter product cycles, volatile markets, and increased demand on customized production.

However, a major obstacle for introducing human–robot collaboration in industry is in how to handle the safety issues. Villani et al. [1] writes that: "Safety issues are the primary main challenge that must be tackled by any approach implementing collaboration between humans and robots". In many robotic automation systems, physical or sensor-based barriers are used to avoid any dangerous situation in human and robot encounters. Collaborative systems, on the other hand, require other measures to ensure safety since robots need to work together in the same physical space as human operators. Many of the models of robots labeled cobots are equipped with safety features such as force sensors and speed limitations and have a CE-marking. When used in an industrial setting, however, they are equipped with tools and combined with other machinery in a way that can create new risks and require new risk assessments.

Regulatory authorities as well as standards organizations have recognized the existence of these smaller robots with safety systems that

* Corresponding author.
E-mail address: atieh.hanna@volvo.com (A. Hanna).

https://doi.org/10.1016/j.rcim.2022.102386
Received 14 July 2021; Received in revised form 18 May 2022; Accepted 22 May 2022
Available online 15 June 2022

are supposed to make them safe to operate in close proximity to humans. There are also standards that are intended to facilitate safe use of collaborative robot solutions. For example, the ISO 10218 and ISO technical specification ISO/TS 15066 [4] entails instructions for how to interpret the Machinery Directive (2006/42/EC) for collaborative applications using the same basic principles for safety evaluation. However, none of these standards take into consideration machines being able to re-plan their routes or make changes under current operations [5] – but rely on the same basic understanding of safety as the regulation of non-collaborative robots and traditional automation.

A traditional automation system mainly has few available actions to choose from at each decision point, i.e. the system has a low degree of freedom (DoF). If something happens unexpectedly or if an action fails, the traditional automation system may stop executing the current task or may even require a repair or a complete reset. It is possible to plan for unexpected events by creating a program that has more alternative actions to choose from. However, to increase the system's degrees of freedom and to make the system truly intelligent and flexible, it must have an online decision algorithm that can plan what task to execute, to handle unforeseen events and to be able to reach its goals.

From the perspective of this paper, a constructive approach to safety of intelligent human robot collaboration is that both humans and robots should be part of the deliberation and planning process, which must be based on negotiations and arriving at consensus decisions [6].[1] This is aligned with Golderberg's [8] preposition for a "constructive and inclusive alternative: 'multiplicity', where humans collaborate with AI and robots to mutually complement each other". His vision summarizes the need for a new intelligent control system that includes both machines and humans into one system. As Malone [9] is stating that the perspective should be changed from "computers in the loop" to "computers in the group" underscoring the importance of maximizing the flexibility and the efficiency of the system by utilizing the specific qualities of humans and computers to create an intelligent solution.

As will be argued in this paper, it is only possible to increase both the flexibility and efficiency when human(s), robot(s) and computer(s) work together in a joint capacity as an intelligent group to effectively reach specific goals in a dynamic environment [9]. This paper studies how such intelligent human–robot collaboration system[2] can be safely implemented in industrial production. We will argue that the required safety precautions necessary for such intelligent systems in some regards are at odds with regulation and safety standards for traditional automation systems that only perform static pre-programmed tasks. From the perspective of this paper, regulation of intelligent automation needs to be approached from a fundamentally different perspective. Rather than focusing on eliminating all risk and adding more restrictions and barriers it should view humans as resources that need to be supported, educated, and trained to deal with the varying conditions of a dynamic human robot interaction (c.f., [10–12])

The overarching aim of this paper is to better understand the safety and regulatory challenges with a focus on the automotive industry, and to suggest a new safety approach called deliberative safety together with a taxonomy of its various safety measures, for intelligent human–robot collaboration. In order to provide a better understanding of the perceived regulatory obstacles to intelligent automation, current views on regulation and safety of human–robot collaboration are studied through interviews conducted with 26 expert stakeholders involved in one regard or the other with the automotive industry in Sweden. The tentative solution to the safety and regulatory problem takes its departure from the problems identified through the interview study, but it also relies on previous research and existing ideas concerning safety and regulation in the scholarly literature.

The paper proceeds as follows. Section 2 presents current relevant regulation and safety standards, as well as current conceptualizations of safety measures in the previous research literature. After that, Section 3 presents the stakeholders views on intelligent collaborative systems in the automotive industry. In Section 4, the new deliberative safety for intelligent human robot collaboration is presented together with a categorization of the possible safety measures that is part of this type of safety. And the final section summarizes the important findings of the paper and discuss some implications of these findings.

2. Current safety standards and regulations and previous conceptualizations of safety measures

The general understanding among the industrial and academic communities, is that the purpose of cobots is to combine the benefits of both humans and robots, to improve quality and productivity [13] and to enable more efficient and flexible assembly lines [14]. Franklin et al. [15] stated that the concrete industrial benefits of cobots are to: 1) safeguard cost reductions, 2) improve space utilization on the factory floor, 3) enabling the redeployment of workers to perform higher-value-added tasks, and to 4) implement partial automation of existing processes.

2.1. Safety of collaborative systems

As discussed in the introduction, Villani et al. [1] understand safety to be one of the main obstacle towards achieving collaboration between humans and robots. The robot's safety standards and the extent to which the standards allow collaboration are also discussed in the survey by Villani et al. [1]. They note that the collaborative modes do not really require a robot that is labeled "collaborative" or "cobot" by the manufacturer, but will also work with ordinary industrial robots, although the safety functions or necessary sensors in such cases needs to be external instead of embedded in the robot. They conclude that "safety issues should be addressed by identifying performance oriented solutions".

Another survey by Saenz et al. [16] states that despite all the potential benefits of cobots, they are far from as wide-spread in the industry as was initially predicted probably due to safety concerns. According to the authors, the ISO safety standards describe how to design, state the requirements for how to protect and how to reduce the risks at a general level, but that each application is unique and when designing a new system, there are many aspects and details to solve. The strict requirements are a great challenge and difficult to meet while considering the business requirements. To address this, Saenz et al. believe that a better design process is needed, that takes into account and comply with the requirements of the standards. In their article on the multimodel interaction between human and robot system Wang et al. [13] stated that "the lack of standards and safety solutions results in a low acceptance of the human–robot combination". They emphasize that the standards play an important role in how safety should be addressed, though they underscore the need of including new technologies and of supplementing required safety measures with mental and emotional safety to make the operators comfortable while performing their tasks.

This is in line with Lasota et al. [17] who defined the safety of human robot collaboration as two folded, the physical safety and the psychological safety. They reviewed the vast body of research focusing on the development of strategies and methods for ensuring safety and summarized them into four main categories: safety through control, motion planning, prediction, and through psychological consideration. Individually, these methods have various benefits and drawbacks, so

[1] The concept of deliberation is commonly used in the social sciences to describe approach to governance [7]. It is commonly associated with the German social philosopher Jürgen Habermas and his theory of deliberative democracy.

[2] Such a collaborative system can include small robots labeled as collaborative (cobots) but they might just as well include other robots for industrial use like large industrial robot arms and autonomous mobile robots.

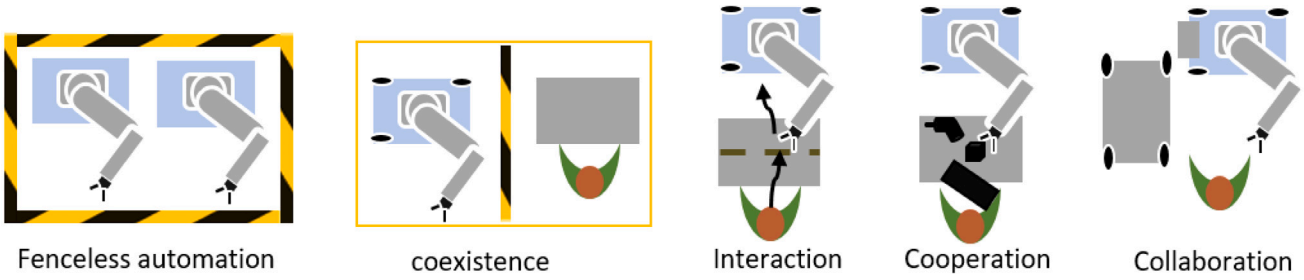


Fig. 1. Human robot relations.

Lasota et al. highlighted the importance to evaluate the combination of several methods to improve the safety for each specific application.

Gualtieri et al. [18] identified two types of safety approaches in the literature, Contact Avoidance, and Contact Detection and Mitigation. To bring intelligent collaborative robotics from the laboratory into the shop floor, they highlight the necessity to further develop and include human-factors approaches that provide methodologies for collaborative systems evaluation and tests. According to Gualtieri et al. this approach could strengthen the concept of human-centered work-spaces supported by smart automation systems.

2.2. Collaboration and interaction

One of the first step when planning a collaborative application, is to identify what type of interaction is required, in order to identify what type of safety system is necessary to comply with the regulations. But there are several different ways of categorizing modes of human–robot relations in the literature, these are commonly referred to as level of collaboration.

While in popular discourse and in marketing the *collaborative* term often simply refers to any robot installation that is "fence-less", the most common taxonomy makes a distinction between: coexistence, interaction, cooperation, and collaboration [13,16] see Fig. 1. There are also taxonomies [19] that makes a distinction between non-inter-active cooperation, sequential cooperation, direct and indirect collaboration. These classifications are based on if humans and robots share space, share time, share task and goal with physical interaction or concurrent task with non contacts [19]. Vicentini [19] reviewed these different taxonomies and provided a comparison and evidence for inconsistencies in the technical literature and in practice. The author underlined that these multiple definitions have scarcely helped in identifying safety requirements nor to comply with safety standards. Instead he proposed a definition of a human–robot application in terms of the frequency of requested access by humans inside a space for shared task execution when both humans and robots are operational and the kinetic energy of robot systems (approximated through velocity). These two terms should, however, be considered in the risk assessment of an application.

2.3. Regulation and risk assessment

In an EU regulatory context the manufacturer, integrator, and end user are required to ensure that the use of the robotic system meet the essential health and safety requirements according to the Machinery Directive (2006/42/EC) [20] either directly through the requirements stipulated in the directive, or using the harmonized standards. The most relevant and applicable standards here are ISO 10218:1 ISO [21] and ISO 10218:2 [22] (type c) as well the technical specification ISO/TS 15066, Robots and robotic devices — Collaborative robots. These standards are voluntary to use but the risk assessment based on the requirements of the Machinery Directive is mandatory for all robotic applications.

The risk assessment is a comprehensive and iterative process that usually engage several stakeholders. It begins with hazards identification, estimation and analysis of their severity and ends up with risk reduction. According to ISO 12100, a risk in a particular dangerous situation, is a function of the severity of the harm that can result in the dangerous situation and the probability of occurrence of that harm. In turn, the probability of occurrence of that harm is defined as a combination of the exposure of a person to the danger, the frequency of the dangerous situation and the opportunity to avoid or to limit the damage.

Protective measures in the standard are the combination of measures to be implemented by the designer of the robotic system as well as by the end user. Risk reduction principles to be implemented by the designer are prioritized according to how effective they are at reducing a risk. In the first phase, risks are tried to be reduced through *inherently safe design*, by eliminating the hazard and its associated risk. If that is not possible, risks could be reduced by providing protection, *safeguarding*, preventing a risk from occurring, or by other *protective measure* to mitigate the risks by limiting possible serious consequences. An alternative risk mitigation is to provide information such at warning signs, signals, instructions, education or similar. Further for residual risk after protective measures implemented by the designer the user should also implement his own protective measures based on the information for use (such as use of personnel protective equipment and training of operators and staff). In addition, ISO 10218-2 provides a list of the most significant hazards associated with robot system to be considered in a risk assessment and describe what potential consequences these hazards can lead to, e.g., mechanical, electrical and ergonomic risk.

2.4. Collaboration modes

There are four methods to ensure safe human–robot collaboration described in the standards ISO10218-2, and in ISO/TS 15066: *Safety-rated Monitored Stop*, *Speed and Separation Monitoring*, *Hand Guiding* and *Power and Force Limiting* (see Fig. 2). *Safety-rated Monitored Stop* is used in applications where the robot system and human need to work side-by-side or need to access the collaborative workspace, but not simultaneously. It is based on the robot system reaching a "protective stop" (Stop category 2 - a controlled stop with power left to the machine actuators) before an operator enters the collaborative area. In this mode the robot remains still until the operator leaves the area then it may resume automatically. Some collaborative robots are featured with motion limitation and force sensors which functions to achieve a protective stop, but also traditional industrial robots can be used with complementary safety devices such as safety cameras or a light curtain that monitor the workspace for any human entrance.

Hand Guiding is used when the operator needs to reprogram the robot paths or reposition the robot. It allows to transmit motion commands to the robot system via a hand-operated device. This mode requires that the robot is equipped with safety-rated stop monitoring to ensure the safety of the operator entering the collaborative area and safety-rated speed monitoring to ensure that the robot is moved at a controlled speed where no additional hazards are introduced.

Speed and Separation Monitoring is used where both operator and robot may need to move concurrently in the collaborative area. The

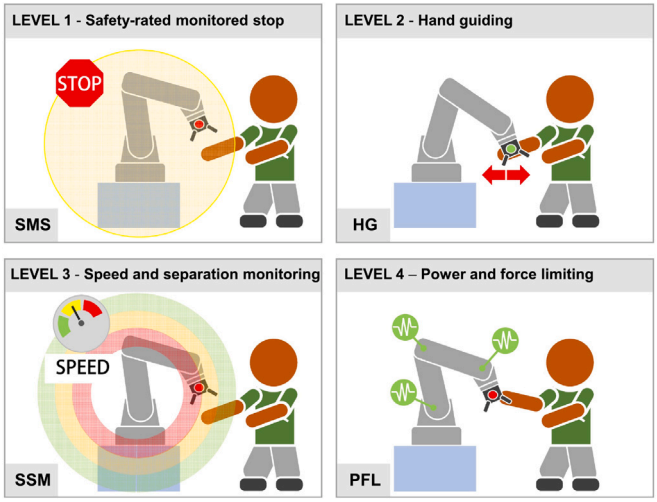


Fig. 2. Collaborative modes according to ISO 10218:201 from Villani et al. [1].

goal of SSM is to anticipate collision between moving resources, the robotic system and humans, by controlling and always maintaining a protective separation distance. During robot motion, the robot system never gets closer to the operator than the protective separation distance. The protective distance, which may change dynamically and in real time, is determined by the robot's speed, robot's reaction time and stop time as well as the human's speed and the safety-rated monitoring system's response time and uncertainties measured by monitoring sensors or robot measurements system. If the separation distance decreases below the protective separation distance, the robot control system may reduce the robot speed followed by a transition to safety-rated monitored stop. Another possibility is to execute an alternative path which maintain the protective separation distance, continuing with active speed and separation monitoring. In this safety mode the robot (or external safety equipment) shall be equipped with safety-rated speed monitoring functions, safety-rated stop monitoring functions, and safety-rated soft-axis, as well as space limiting functions.

Power and Force Limiting is used when the robot and the operator can have physical contact intentionally or unintentionally. It requires the robot system to have a passive and inherently safe design such as smooth surfaces and that the robot (or external safety equipment) is equipped with an active safety-related control system such as sensitive force monitoring functions that can dissipate collision forces and that do not violate determined force limits. Detailed information about guidance on force, speed, separation distance, bio-mechanical and contact limits and how to measure these can be found in ISO/TS 15066.

2.5. Flexibility and collaboration

Much of the scholarly articles on the use and safety of collaborative automation recognizes that applications in the industry have a relatively low degree of complexity in the interaction between human and machine, i.e., they are mostly used as a tool. Several of these surveys understand that more autonomy, intelligence, and cognitive processing skills would be beneficial to be able to utilize them to their fullest capacity. However, in previous work [5], the first author of this paper has showed that the robotic safety standards are limited in supporting the implementation of intelligence and autonomy (to enable efficient and flexible automation solutions), and that they do not consider the human factors incorporating the cognitive and perceptual factors, and do not support the trade-off between the various industrial requirements and safety which result in inefficient and inflexible automation solutions [5].

Gualtieri et al. [23] compiled design guidelines to support designer to develop and evaluate safe, human-centered and efficient collaborative assembly workstations. For Physical safety and physical ergonomics the start point was from safety and ergonomics standards while for the cognitive safety and efficiency aspect they relied on the explanation in research work due to lack of standards cover these perspectives. Another lack identified by Gervasi et al. [24] is the lack of a general framework to evaluate HRC application that consider all aspects of the interaction between human and robots including latent dimensions such as autonomy, adaptivity and training as well as human factors to mention some. The authors proposed different evaluation scale for these dimensions to be considered and shaped by the application designer.

While much of the existing literature on *flexible* automation concerns coordination of machines, there are also several articles using the concept to discuss human–robot interaction. Miller and Parasuraman [25] argue that "Operators need to be able to delegate tasks to automation, and to receive performance feedback, in much the same way that delegation is performed in successful human–human organizations: at various levels of detail and granularity and with various constraints, stipulations, contingencies, and alternatives". Miller and Parasuraman suggest that the efficiency in production could be increased by continually changing the Level of Automation (LOA) to be well suited to the task at hand. Further they write that a perceived problem with an approach of continues changes to the LOA is the unwillingness among user to accept adaptive changes that might be understood as increasing unpredictability of the system due to the fact that humans are not in control of all the actions of the automated system. As a response to this they propose a "delegation approach to adaptable human-automation interaction" so that humans remain in charge and can decide how much automation to use where the delegation is a process of assigning specific roles and responsibilities for the tasks at hand [25].

To conclude, previous research both highlight and discuss safety issues in human–robot collaboration, promote for more intelligence and autonomy as well as propose concepts of human–human like collaboration. Also, concerns with safety standards and the need of updating them are recognized. However, there is need to explore the current perceived understanding of the users and other stakeholders on the intelligent human–robot collaboration which is lacking in the scholarly literature.

3. Stakeholders' view on intelligent human robot collaboration

To better understand stakeholders' perspectives on collaborative robots and intelligent human robot collaboration interviews have been made with 26 societal stakeholders and were collected between January 2019 and March 2020. These stakeholders are engaged in one way or the other with automation within the automotive industry, and work in industry (n=13), university (n=8), government agencies, and other organizations (n=5). The selection of interviewed stakeholders is made to include various interests as stakeholders are known to have different positions and attitudes in regards to development and regulation [26].

Most of the industry stakeholders worked within a large company within the automotive industry (n=11) while n=2 interviewees are from smaller sub-contractors working partly with the automotive industry as customers. All respondents have experience of working with collaborative robots in some regard, and together, these experts have a vast experience in working with issues of relevance to the inquiry of this report. The focus in this article is the perspective of the industry stakeholders and the perspective of the established automotive industry. These interviewed stakeholders are generally well-educated with 9 of them holding a PhD degree. In regard to gender division n=21 were male and n=5 were female, reflecting a strong male dominance within engineering [27]. For a more in-depth description of the perspectives of these stakeholders see the report [28].

The interviews combined a semi-structured format with open conversations where the respondents were invited to speak freely around the topics of inquiry of this paper. In the semi-structured component, some general key questions were asked in the same way in all interviews. These key questions was about their background and organizational role, their views on benefits and obstacle towards introducing intelligent automation in the automotive industry, their view on the future of automation and the need of humans in production, as well as their views on regulatory obstacles. The open approach was intended to capture each participant's understanding, and to encourage reasoning, reactions and own considerations on the topic [29,30]. The interviews were adjusted to the different categories of interviewees, government authorities were mainly asked about their perceived role in this technological development, industry and researchers were asked more specific questions regarding their view on the specific technologies. The interviews lasted between 38 and 200 min, with an average of about 90 min.

All interviews were transcribed and then inductively coded. The coding of the transcribed material into thematic categories was both open and selective, in line with the basic principles of grounded theory [31]. This qualitative approach addresses data collection and analysis as interrelated activities. It is based on thematic coding from the basis of common themes in the interview material. The interview material was read a number of times and codes were formulated through interpretation and comparison within and between interviews [32].

3.1. The perceived benefits

In these interviews potential benefits of human robot collaboration in the factories are stressed by most respondents, especially researchers and those industry stakeholders working with research and development. There is a general understanding that collaborative solutions potentially could improve the production in final assembly that today is done mainly manually. Envisioned improvement among the industry respondents include incremental improvements in ergonomics, quality, and safety of production by "removing work tasks that are potentially harmful to the operator". Safety benefits are especially stressed when talking about replacing manually operated trucks and forklifts with automated guided carts, but also collaborative solutions using cobots are understood as solutions to further improve safety. There are also suggestions of uses that could increase flexibility in the production. For example, one presented idea is that a fleet of small automated guided carts (AGC) with a central control system coordinating the actions of carts could provide a more streamlined and safer way of transporting goods from pre-assembly to the main production line. Several of the respondents also stress potential benefits using mobile cobots that could be moved between workstations to balance the flow in the production due to variability in products. Other ideas of how to use intelligent automation include cognitive support for operators.

3.2. The middle management

While all respondents talk about potential benefits, industry stakeholders in middle management working with technology upgrades in the production of final assembly remain largely reluctant to collaborative intelligent automation. It is considered that these solutions are difficult to design but that they can add value to the production. More specifically, the stated reasons for the difficulties in finding good use for intelligent collaborative applications are that available equipment for smart automation such as cobots and automated guided carts are not sufficiently developed for industrial use, and that the products that they manufacture are not designed for automation. Furthermore, it is considered to be difficult to introduce intelligent automation in existing processes, relating both to incompatibility between new and old technology as well as a lack of compatibility with current routines and established ways of doing things in the factory. It is generally not considered impossible to design automation with human robot collaboration but the available technology and equipment is understood to makes these solutions slow and expensive. For example one of the interviewed industry stakeholders said: "In theory the robot could find a screw head using these new technologies, the problem is that it is very time consuming [...] and time is money".

3.3. Pilot projects and bad experience

When talking about obstacles to introducing collaborative automation the industry, stakeholders often refer to pilot projects in the factories, where a cobot has been installed saying that the installation functions poorly or that the costs of the installation have exceeded the benefits. Rather than seeing immediate benefits of these pilot projects most respondents say that they do such projects to evaluate these emerging technologies. A stakeholder says: "you could say that we are exploring the technology, but we do not implement it [on a large scale] if we do not see a profitability". Although there are some installations using cobots in the factories, one of the interviewees state that "when we install a collaborative robot it rarely does any collaborative work [...] we do not have to fence it, but it is very rare for us to find applications were humans and machines can collaborate". So rather than doing collaborative work, the cobots are programmed and used in the same fashion as traditional robots in the industry. Consequently, when criticizing collaborative intelligent applications the industry stakeholders are not really evaluating the potential functionality of intelligent human robot collaboration but rather the functionality of "cobots" as such.

So, while the interviewed industry stakeholders recognize the benefits of a flexible production facing a future with shorter product cycles and an increased diversification of products they do not understand collaborative automation as a response to this. In stark contrast to the ambition of manufacturers of cobots that collaborative automation could make production more flexible, from the perspective of the industry stakeholders, human–robot collaborative solutions are not flexible or helpful in dealing with large product variability. In fact, large product variability is understood as the largest obstacle towards introducing automation in the production. An industry stakeholder say: "In order for it to be possible to do it at all, there needs to be a fairly low variation between the products, otherwise it will be very difficult to tell the robot what to do". When discussing possible areas to automate with cobots and AGC's in the factory, the industry stakeholder stresses that this is only possible in segments with low product variability and repetitive work tasks.

3.4. The challenges with current safety regulation

The safety of intelligent automation is regulated in the same way as any machines or robots in the factory according to the Machinery Directive. All machines and machine lines have to be CE-marked, meaning that the manufacturer signs the Declaration of Conformity after making a thorough risk assessment and documentation. The process of making this risk assessment for intelligent automation and collaborative robots is in one regard very similar to the evaluation of any other technology. An operator at Volvo GTO says: "We have a protocol for risk assessment [...], by which we assess all the work tasks in here, there is really no difference [with intelligent automation] compared to other tools and robots that we install". So, while intelligent automation is being evaluated with the same principles as other robots and equipment in the factories, some things are different, especially in regard to the introduction of intelligent automation operating in close proximity to humans. The understanding among the industry stakeholders is that any robots in areas that are not fenced off creates problems in relation to current regulation.

The smaller robots (cobots) that can be used in collaborative stations are in most cases CE-marked by the manufacturer according to

european product directive such as 2014/35/EU — Low Voltage Directive (LVD), 2014/30/EU — Electromagnetic Compatibility Directive (EMC), 2011/65/-EU — Restriction of the use of certain hazardous substances (RoHS). The robot itself is "partly completed machinery" and therefore is not safe but often has internal certified safety features. The robot must be part of a CE marked machine before it can be used to comply with the Essential Health and Safety Requirements (EHSRs) provided by the Machinery Directive. When applied to a new setting in the factory the safety has to be evaluated through final risk assessment and the application CE-marked, and someone must be listed as the "integrator" for the installation. An industry stakeholder say that "it does not matter if the cobot has [built-in safety systems], you still have to do a risk assessment of the overall system". Another stakeholder explains: "We can buy a cobot today, which is CE marked, but if we equip it with a knife, it is no longer safe". Furthermore, risk assessments become more difficult when robots become smarter and re-plan routes according to outer stimuli, and optimize their actions due to previous experiences.

The main problem with the regulation of collaborative automation from the perspective of the stakeholders is that the robots do not operate within fenced areas but in proximity to humans, and that the robot might act in ways that is (or is perceived as) autonomous which makes the risk evaluation more difficult. The current regulation is considered to be too restrictive in enabling efficient intelligent solutions in the factory – while safety is also one of the most cited reasons to introduce such automation – in other words current regulation is not aligned with existing norms In the wordings of the industry stakeholders current regulation is not adjusted to what is understood as "tolerable risk" and "common sense".

When talking about regulatory obstacles, the stakeholders general understanding is not that it makes collaborative solutions impossible to implement, rather that the process for CE-marking, and the safety directives require redundant safety systems that make the solutions slow and inefficient. An industry stakeholder say:

> The big challenge will be to resolve the safety issues in a way that is reasonably cost-effective, you can always build any number of safety systems and make it safe, but then you get a very rigid and expensive installation. [...] And also, we have to solve the safety issues in a way so that the robot installation becomes efficient. You can always resolve the safety concerns by making the robot go very, very slowly, but then the robot will be useless.

In the view of the industry stakeholder the regulatory requirements make smart automated solutions too slow and it makes it difficult to find useful applications for cobots. Furthermore, every time the robots are used in a different fashion (in another location or equipped with other tools) a new risk evaluation has to be made and the documentation for the CE-marking needs to be updated—which is understood as a further obstacle towards flexible use of intelligent automation systems. The requirement to make new CE-markings or update the documentation for every new application in the factory is seen as making it too complicated and expensive to achieve desired flexibility. Current regulation is thus understood to be an obstacle for a flexible use of smart technology. So, in contrast to what many manufacturers of "cobots" argue, the possibility to use the robots for several different tasks in the factory is not understood as viable by industry stakeholders within the automotive industry.

3.5. Accepting some risks

The stakeholders that saw current regulation as an obstacle towards introducing intelligent and collaborative automation were asked if they had any suggestions for changes that would facilitate large scale implementation of collaborative solutions. The respondents did not provide any ready-made solutions for how such new regulation should be designed. However, there are suggestions that risk assessment should not only rely on physical barriers but rather on education, communication, and trust. From such a perspective the respondents also understand it as necessary to make a certain degree of risk acceptable: "if the operators get an education in these systems, they will consider these risks as acceptable". Some degree of risk acceptance is considered necessary to accept in order to facilitate what they consider to be safer solutions compared to current production (thus recognizing a need of including risk trade offs in safety evaluations). There are also suggestions of regulating intelligent automation as animate objects rather than machines, for example by viewing and regulating smart robots in the same way as livestock as a way of recognizing the contingent behavior of intelligent robots. Another idea discussed by one of the industry stakeholders, is to make dynamic solutions that could be adjusted to variations in work mode: "We are looking at solutions where a collaborative robot senses when people are nearby and adapts to it, but it can also operate at high speeds and use enough force to handle the weights of our products, it is interesting to be able to operate a robot at high speed, but when humans approach it slows down and you can work together". Due to these perceived difficulties the interviewed stakeholders do not understand cobots to be a flexible solution or a solution that could help make the production more resilient, quite to the contrary from how this technology is discussed in the marketing of cobots, and in some of the previous research.

The views of the stakeholders can be summarized in the following bullets:

- Stakeholders see potential benefits in intelligent collaborative solutions to help automate parts of the final assembly.
- Intelligent solutions in the production are considered to be more safe than current manufacturing.

However, benefits associated with collaborative automation are understood to be difficult to achieve due to:

- Cobots are understood to have a low degree of dependability compromising the potential benefits of streamlining the production.
- Current available collaborative solutions/cobots are not considered to be a cost effective solution.
- Cobots are not considered to be a flexible solution.
- Current regulation is not understood to be well adapted to robots operating in the same physical space as humans.
- Current regulation is understood to make human robot collaboration inflexible and slow by requiring redundant safety measures.

While manufacturers of collaborative robots as well as research literature stress the benefits of intelligent collaborative automation in streamlining the production as well as making it more flexible they rarely offer convincing arguments as to how this should be achieved. The difficulties in designing intelligent collaborative automation that can add value to the final assembly is confirmed with the interviews with industry stakeholders presented above. An important limitation to achieving benefits is that cobots mainly are used as traditional robots, and their performance is evaluated in the same way as traditional robots that have a larger capacity. Furthermore, finding more innovative uses is in many regards restricted by current regulatory framework. In the section that follows we will suggest a framework for addressing safety issues in human collaborative automation addressing the identified concerns of the interviewed stakeholders based on current regulation.

4. Deliberative safety and its safety measures for intelligent human robot collaboration

Many stakeholders have the experience that most cobot installations are inflexible, inefficient, and not very robust. This is due to a general confusion about the difference between traditional and intelligent

automation, resulting in industry stakeholders failing to see potential benefits of collaborative automation. For a human–robot collaborative system to be both flexible and efficient, the system must be highly adaptive, be able to cope with unforeseen events, handle variations of parts and resources, be able to interact with other humans and robots, and have a large degree of freedom to perform complex operations. If such flexibility can be achieved while also efficiently accomplishing various goals, it can be defined as an intelligent human–robot collaboration system. However, this fashion of intelligent human–robot collaboration requires different kinds of safety measures compared to traditional automation before it can be installed in real industrial environments. The required safety measure of a system is also depending on the required level of flexibility and efficiency, as well as the budget for the installation in question.

To better support the introduction of intelligent human–robot collaboration, we suggest a new safety approach called deliberative safety, where the current safety measure can be changed during operation based on the need and intention of both the humans and robots in the system. Either simple measures can be used such as creating visual barriers, defining speed limiting zones or force and torque limitations, or more advanced measures such as active or planned safety. But before studying this new approach and its measures, let us define a number of performance indicators to differentiate the various measures.

4.1. Performance indicators of safety measures

In an intelligent system where robots and humans together try to accomplish a goal, the work can be divided into multiple operations, where an operation consists of a number of small tasks. If we for example have a system that is picking parts in a warehouse, doing some minor assembly, and then transporting the kit of parts out to another system, there will be operations where humans and robots collaborate and others where each resource is working individually. Some operations are not related to the ‘‘production’’, such as recovering from failures, troubleshooting, learning new tasks, giving or getting instructions, handling variations or planning what to do. Also for these types of operations, all resources in the system need to collaborate and potential risky situations may occur.

When deciding what type of safety measures that are needed for each operation, the required measures will be dependent on many different requirements and may even vary for the same operation over time. In this paper, deliberative safety defines a taxonomy of possible safety measures, so to be able to define them, evaluate them and decide which measure to use for an operation, a number of key performance indicators (KPI) are defined in Table 1, namely the four flexibility indicators required by an operation: *Degrees of Freedom*, *Interaction*, *Speed*, *Concurrency*. There is also the three implementation indicators *Coordination*, *Monitoring*, and *Complexity*, which defines the challenges when implementing a measure.

The four flexibility indicators are used for categorizing an operation and what requirement the overall performance of the system has on that operation. If the implementation of a specific operation must have a high degree of freedom, it is important that also the safety measure that is used will allow that. When combining the various flexibility KPIs of an operation, some measures will be better to use than others.

The input for the creation of these KPI were based on experience from several installations done both by a focus group and the research group but also based on the literature and the lack of safety standards covering the aspects of efficiency and flexibility. To start with, we suggested a simplified scale from low to high for these KPI to be able to trade-off between the different performance and flexibility requirements. In the future this scale may need to be refined and weighted after evaluation on a number of cases.

One way to implement a safety measure is to restrict the number of possible actions available to the humans and robots, by for example not allowing any interaction or forcing a fixed pre-planned sequence, i.e. having low *Degrees of Freedom* (DOF). However, when a system is allowed by the safety measure to have a high DOF, both the humans and robots must adapt their behavior to handle uncertainty and changes in the environment. Another indication if a safety measure allows a flexible control is the level of physical *Interaction*, since different measures will allow a varying degree of physical interaction. The level of required physical interaction between humans and robots during an operation is important to determine before deciding on the relevant safety measure. For example, if they only will be working side by side or in sequence, the humans and robots are having less physical interaction, or if the human is hand guiding the robot while teaching a new task, more interaction is needed.

It is sometimes possible to increase the efficiency of a system by moving as fast as possible, making the possible *Speed* of the robots an important indicator. The faster they can move, the shorter the execution time (in most cases), but it also entails an increased risk of accidents if it is operating in the same space as humans. Another way to increase the efficiency of a system is to execute as many tasks *concurrently* of an operation as possible. For some operations, this is only feasible if the resources can work at the same time in the same area.

Some safety measures will allow an increased flexibility of the system but may also entail an increase in investment in both development and in training of operators since they increase the *Complexity*. The measure can for example require a complex control development, code certification, extensive testing, calibration or training. Sometimes, the gain in flexibility will not motivate an increase in investment. Over time, the investment required for complex safety measures may decrease when they become more common, but they may always require a larger investment than simpler safety measures. One part of this complexity comes from a need for the humans and robots to coordinate their actions and communicate their intent. If a robust *Coordination* is possible, some safety measures will then allow both flexibility as well as efficient execution. Another important indicator that may increase the complexity of a safety measure is how the resources and humans are *monitored*. Sometimes advanced vision and AI based monitoring is required and sometimes it is enough with some buttons or a light curtain. The type of monitoring needed will influence the required investment, but if for example an advanced monitoring system is used, it can increase both the flexibility and efficiency when combined with an intelligent control system that can utilize the information.

4.2. Deliberative safety

The overall aim of this paper is to enable intelligent human–robot collaborative systems in real industrial installations in which humans and robots can sustainably work together in a safe way that is flexible, adaptive and efficient. However, the challenge is that current standardized safety measure do not allow flexible and intelligent control systems, and some newer types of safety measures (e.g. [18]) are not very efficient.

Therefore, we are proposing deliberative safety, which allow the humans and robots to switch between different safety measure based on the need for flexibility or efficiency. To make the performance of an operation cost effective it might sometimes be necessary to allow a high level of flexibility and at other times be able to be executed as fast as possible. By allowing the humans and robots to decide about the required performance level, while they are safely performing the operation, a highly flexible yet efficient system can be created.

The idea of adaptive automation of tasks for streamlining and improving automation has been around for quite some time e.g. [25, 33–35]. The idea in these literature is that levels of automation are beneficial for different tasks and that this should be adjustable in an efficient manufacturing system. The proposed deliberative safety instead stresses the importance of adaptive safety measures as a way of addressing both the issue of performance and the safety issue. By

Table 1
Key Performance indicators for evaluating the need of flexibility for an operation and the implementation challenges when implementing different safety measures. The scale is from low to high. Currently, there are no discrete scale between these two levels, but in the future, based on usage experience of these KPIs, new levels may be added.

Flexibility indicator	Definition
Degrees of Freedom (DoF)	Defines the required degrees of freedom where a high DoF means that the operation must be highly flexible and adaptive in the sense that it can re-plan the sequence of the operation or change actions depending on surrounding events. A low degree of freedom means that the operation only need a fixed pre-planned sequence with predefined alternatives to reach its goal.
Interaction	Defines the required interaction between the humans and robots. To successfully perform the operation, high interaction means that the robots and the humans must interact in some form, with e.g. physical contact, and to give and receive feedback. A low degree of interaction means that the humans and the robots can perform their respective tasks independently without interaction (or if only one resource is used during the operation).
Speed	Defines the required speed of the robots and other resources. A high speed KPI mean that a specific operation must be performed using high speed to be efficient and useful. A low speed means that the operation is not time critical or there are other requirement from the environment to lower the speed e.g.to enable physical interaction.
Concurrency	Defines if multiple humans and robots must work in the same area at the same time, where a high concurrency KPI means that resources must work concurrently in the same area to be able to perform the operation. A low concurrency KPI means that the resources do not need to work in the same workspace at the same time.
Implementation indicator	**Definition**
Coordination	Defines the required level of coordination used in the safety measure, where a high KPI means that the humans and robots must robustly coordinate and communicate among each other to stay safe. A low KPI means that no coordination is necessary to stay safe.
Monitoring	Defines the requirements a specific safety measure has on the monitoring system, where a high KPI means that the safety measure requires that the monitoring system must include advanced functions (e.g. using vision or predictive algorithms). If a safety measure has low requirements on the monitoring system, it means that a very simple monitoring (e.g. a fence) or no monitoring is required.
Complexity	Defines how complex a measure is to implement, where a higher KPI means a more complex solution and often a higher investment. A low complexity means that the measure is often simple to implement.

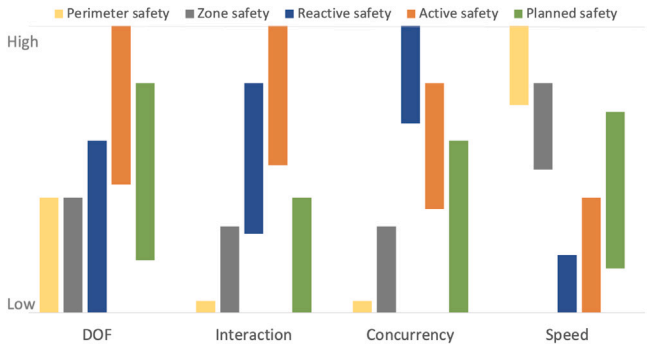


Fig. 3. Each safety measure is suitable to use for different types of operation requirements. The suitable range of a flexibility KPI for each measure is shown in this figure and can be used for deciding what measures to use for each operation.

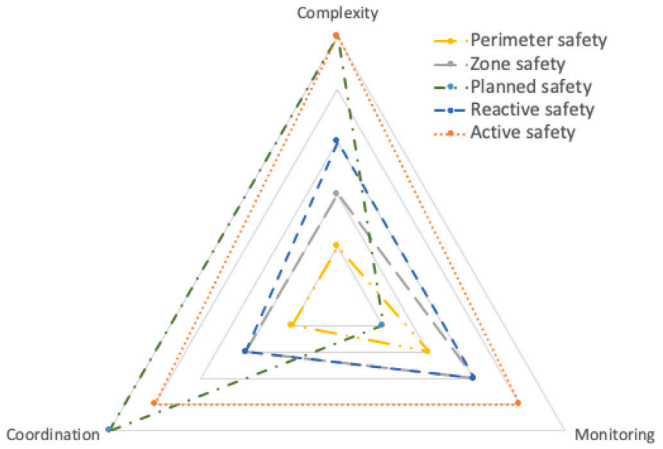


Fig. 4. The three implementation KPI's for the five safety measures used in deliberative safety. Perimeter safety, zone safety and reactive safety are quite easy to implement and active and planned safety are harder to implement.

drawing on the idea of continuous changes in safety measure to suit the tasks at hand, this will not only facilitate more efficient automation solutions but that it also can be used as a safety measure for intelligent robots and machines in production.

When deliberative safety is used, the system can switch between various safety measures. To better evaluate these measures for deliberative safety, possible safety measures are discussed below and their performance indicators are shown in the two charts Figs. 3 and 4.

Perimeter safety

This safety measure is sometimes called "Safety Monitored Stop" and safety risks are avoided by not allowing the robots to operate when humans are in their working areas. This can be achieved for example with sensors or light curtains. The benefit of this safety measure is that the robot can move very fast when the human operator is outside the work space of the robot and the safety measure is easy to implement both from a control and monitoring perspective. However, this safety measure limits the DoF, and does not allow any interaction or concurrency of an operation. This measure is commonly used in industry today where robots and humans need to work in the same physical space but not simultaneously.

This measure is suitable to use when really high speeds are necessary and if a secure perimeter monitoring can be used. It is in many cases the preferred measure for the operations where a robot is working by itself. It is not very complex to implement, the monitoring system is based on well known technology and the required coordination and deliberation between the human and the robot is simple.

Zone safety

If somewhat more flexibility is needed compared to perimeter safety, the work area of the robot can be divided into multiple zones, where the speed and behavior of the robot is restricted based on what zones humans or other resources are present in. This is sometimes referred to as "speed and separation monitoring" and is often used

to reduce the speed of the robot when a human is approaching, but the robot could also change any type of behavior to increase safety. This safety measure requires a monitoring that can divide and monitor multiple areas in the station, and is more complicated to install and program compared to perimeter safety but still a relatively low degree of complexity compared to active safety (see below). This measure is also common to use in industry, but it require space and may limit the efficiency when humans and robots need to work together.

This measure is also mainly used when speed is necessary and when their is no need for a high DoF, but zone safety will allow a little bit more interaction or concurrency in the operation compared to perimeter safety. The trade-off is that it is more complex to implement and a more advanced monitoring and control system is required.

Reactive safety

A completely different approach to create safe motions is to only move the robot slowly and monitor the forces of the robot to keep the potential impact within the requirements stipulated in the Machinery Directive. This is sometimes called "power and force limiting", or Contact Detection and Mitigation [18], and does not require any external monitoring other than on the robot. This is a common safety measure for cobot application today and it makes it possible for the robot and operator to work together and to interact. However, it requires the robot to move very slowly and since the system is unaware of the human, it is complicated to include an intelligent control strategy.

The most important benefit of reactive safety is that it allows a very high degree of concurrency since all resources can just keep on working without any safety risks. The monitoring is simple and there is no need for complex coordination. But the possible DoF is not very high and the interaction among the resources are limited to reactive actions. This measure will mainly be used in classical cobot applications where the robots and humans are coexisting in the same work space and the efficiency of the operation is not of importance.

Active safety

By combining advanced monitoring with the use of e.g. vision and predictive algorithms, it is possible to avoid collisions by for example actively changing the robot trajectories or the sequence [36], using tools for assisting the human operator [37], or using advanced sensors [38]. To create a truly intelligent collaborative system, active safety is often necessary since it allows great flexibility and possibilities for interaction and communication. Since the monitoring system can in some cases keep track of the human movements and predict human intentions, it is also possible to run the robot faster in some situations compared to reactive safety.

Active safety is necessary when an operation requires a high level of DoF and interaction where the human and robot will together complete tasks and the robot can adjust and adopt its motions based on what the human is doing. Active safety is however not as good in concurrency compared to reactive safety due to the more complex monitoring system. The complexity of the control system, monitoring system and the coordination among the resources, will also limit the applicability of this measure since it may be hard to always guarantee safe behavior This measure is still an active research area [18].

Planned safety

Perimeter, zone and reactive safety are designed to be safe no matter what the human does and requires monitoring of the work area using e.g. safety scanners or just move very slow. Active safety allows great DoF, flexibility and interactions, but it is often challenging to implement and requires a great investment. Another problem with active safety is that the robot will be autonomous and may change its behavior during its execution, which may make the human operator uncertain, not feel safe and may result in a risky situation. Another approach to guarantee safety is instead to let the robot and human plan the sequence of tasks they will perform, and then follow that plan. If a human knows exactly what the robot will do and where to move, it is possible to be closer to the robot than if the robot can change its path.

The planned safety measure includes a well defined interaction point before the execution is starting where the robot and human is deciding exactly what to do and in what order (Exactly how this deliberation will take place is still an active research topic, but will involve some kind of human–machine interaction). When it is decided, they start performing their plans. When something unexpected happens, the work is stopped and they replan before continuing. This allows for a higher DoF compared to perimeter, zone and reactive safety, and allows for some concurrency. It can also allow somewhat higher speeds compared to reactive and active safety, but it does not allow very much interaction during the execution of the plan.

Planned safety is complex to implement and requires a well defined coordination between the human and robot, but the monitoring system does not need to be as complex as active safety. This measure will probably mainly be used together with other measures when using deliberative safety since it already include the deliberation step.

4.3. The deliberation process: a use case

An important comment to make is that the safety of a system that uses deliberative safety is always equally safe, no matter what safety measure that is currently used. When designing a system, a detailed risk assessment and evaluation of each operation has to be made [5] and based on this, the possible safety measures that should be included is selected for each operation. Some operations may only allow a single measure and others may allow the system to decide.

To evaluate deliberative safety, we have implemented an industrial use case of collaborative kitting, where a part kit is prepared in a warehouse and transported to the point of use on autonomous transport robots(ATR), see Fig. 5. The starting point for designing such system is to consider product and production requirements e.g variation of the products and the environments, cycle time and available space. These requirements together with the main goal of the automation, namely transport the kit in time determine how the tasks are allocated to operator and robots. Takatas et al. [39], Dianatfar et al. [40] developed approaches for task allocation taking into account performance, ergonomic and safety criteria which were appropriate for predefined task set-up.

In our use case, the main motivation is to free up operator time by eliminating non added value tasks and ergonomic saving by eliminating heavy lift and be able to reach the production goal in different ways. Therefore it is convenient to let the robot pick the appropriate items from dedicated area in the kitting facade or pallet, place them on the kitting transporter (see Fig. 6) while the operator pick some other items. They will collaborate for example to pre-assemble sub-assemblies, as well as support each other and switch tasks e.g. operator pick items when the robot fails or robot pick items when operator needs to perform other tasks. During the design phase, a sequence is not prepared, but rather any task that the robot can perform is implemented, which is dependent on its capabilities. During the execution of the system, an online planner and optimizer decides who should perform what task based on the state of the system. An evaluation of some included operations according to the flexibility KPIs is given in Table 2.

Some operations e.g. picking and placing items can be done by the robot without the operator's intervention while the operator does other operations. Here the requirements on *DoF, interaction* and *concurrency* are low while medium on the *speed*. Then they can agree to run "zone safety" where the robot is allowed to run at higher speed when the operator is not close by. The safety zones are monitored and the speed is reduces when the operator is getting too close.

A probable hazard can be that the robot fails to pick or loose the grip and the part falls and then the operator may want to handle the error by picking the part. For this operation, the requirement may be

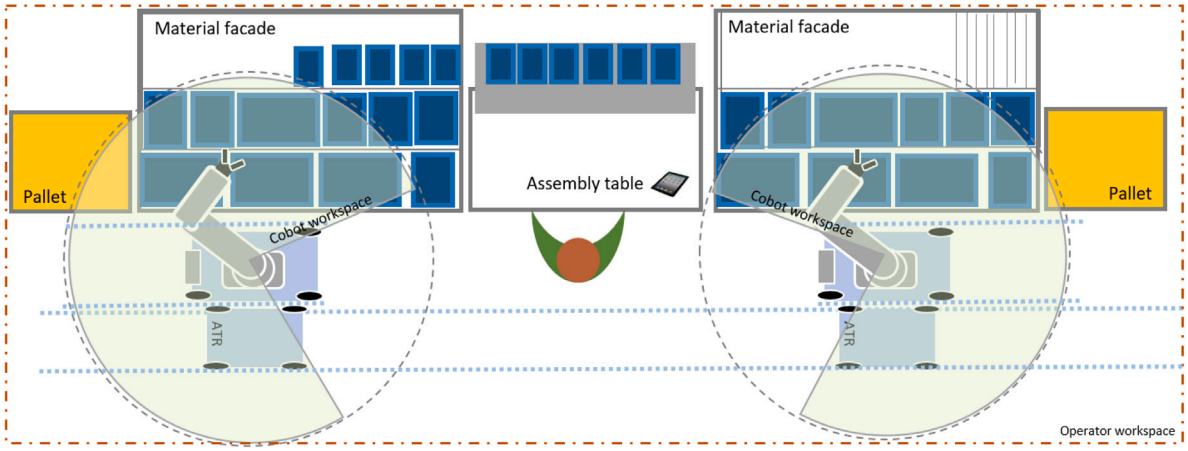


Fig. 5. Collaborative kitting.

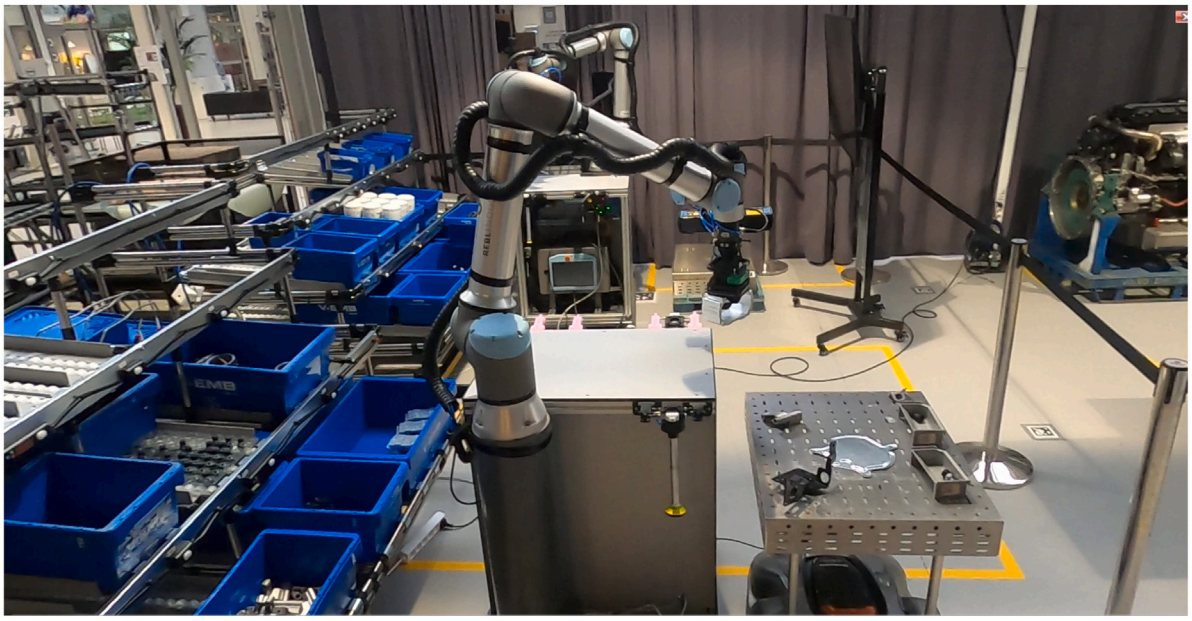


Fig. 6. Use case implementation in a lab environment.

medium on *DoF* and *interaction* while high on *concurrency* and low on *speed*. Instead of stopping its execution, the robot will enter its "reactive safety" mode where it continues to work with a lowered speed, without any risk to the operator. When the operator have supported the robot and moves out of its safety zone, it will switch back to the "zone safety" measure.

In another situation, the robot is delayed and will not have the time to pick all parts for the part kit. Then the robot enters its planned safety measure and will signals to the operator that it needs help. When the operator arrives, the robot suggest by highlighting using lamps (Pick-to-Light system), on which side of the facade it will pick parts and where the operator should pick. The operators agree by pressing a button and they start working. This operation require high level for *DoF, interaction* and *concurrency* but may need medium level of *speed* to enable higher level of efficiency. Even if the operator is in the normal robot safety zone (on one side of the robot), they both can work at medium speed since they will not share working area. The operator may potentially climb onto the robot and be hit, but this kind of risks should be mitigated by fostering a human behavior built on responsibility to respect the agreed plan and on knowledge and education. This will require high level of coordination and communication to ensure the start and end of this measure and if something go wrong the operator have the mandate to stop the execution. With this measure higher efficiency may be enabled with minimum limitation in comparison with perimeter and zone safety but also with less integration complexity in comparison with active safety.

Other operations such as a pre-assembly, where the robot and the operator need to work close, sequentially share the tasks or to coordinate their moves. With other word, this operation may require high level of *DoF*, *interaction* and *concurrency* while it is important to have low *speed*. Then they can reason and deliberate, if possible, to run "active safety". A vision system and prediction algorithm can provide information about the current situation and what will happen next, track humans and signal if something goes wrong or goes not as planned. Then the robot can run a bit slower to have time to react and act in the right way while being able to re-plan if needed. In such operation, a major risk is to collide unintentionally due to system or human failure. In order to mitigate such risks an intuitive human–robot interface, enabling clear communication, is necessary.

For the last operation, transporting the kitted parts from the kitting area to the main line, low level of *DoF, interaction* and *concurrency* is required while high level of *speed* is needed. The operator and the

Table 2
Evaluation of a number of operations from the use case according to the flexibility KPIs.

Operations	KPI	Evaluation	Safety measure
Picking and placing by robot	DoF	Low	
	Interaction	Low	
	Concurrency	Low	
	Speed	Medium	Zone safety
Operator handles delays	DoF	High	
	Interaction	High	
	Concurrency	High	
	Speed	Medium	Planned safety
Operator handles error	DoF	Medium	
	Interaction	Medium	
	Concurrency	High	
	Speed	Low	Reactive safety
Pre-assembly	DoF	High	
	Interaction	High	
	Concurrency	High	
	Speed	Low	Active safety
Navigation	DoF	Low	
	Interaction	Low	
	Concurrency	Low	
	Speed	High	Perimeter safety

system can switch again if it is possible to run as pre-planned. Since no intervention of human is needed and the main risk is to collide with obstacles or humans, a simpler measure can be used such as "perimeter safety". Being able to shift between safety measures allows for highly adaptive automation solutions, however, the operator should have the full mandate in all safety measures to decide to stop the execution if it is necessary.

To be able to achieve this type of flexibility and efficiency in a production system yet safe as we described in our use case, it is necessary to be able to handle unexpected changes. It can be impossible to achieve this with traditional automation, which has usually very few alternatives to fulfill its goal, explicitly programmed for certain sequences (with if, then, else), have predefined layout (mostly fixed), and predetermined safety measures limiting collaboration and deliberation. In traditional implementation, to handle errors such as robot fails to pick or looses grip of a part, may require a resetting of the system to an initial state. It is not possible to handle delays and replan a task from operator to robot or vice versa. On the other hand the aim with intelligent automation that can be adaptive to changes in the environment by learning and predicting, can replan paths or sequences and enable deliberation in planning and acting safely, is to handle such unexpected changes.

Certainly this flexibility in the environment and the sequence may arise many challenges linked to how different dimensions and speeds should be calculated not least how to ensure the delivery on time. Therefore a risk assessment which is mandatory in all automation applications must be performed in several iteration, from the planning phase to final testing for a system. If there are risks with changing the safety measure or mitigation of the risks is only possible by maintaining distance, speed or power and force limiting then the standard's recommended limits should be followed for "perimeter safety", "zone safety" and "reactive safety".

5. Concluding discussion

The approach in this paper does not view (level of) intelligence as a mode of automation, but it refers to the adaptability of the system itself. Thus, the level of automation is not related to the degrees of freedom of the components of the system, but rather the capability of the system to optimize speed, degrees of freedom, and safety to correspond to the particular task at hand. In other words, the level of intelligence is dependent on the automation system's capacity to delegate tasks and set the correct safety measure to optimize the performance of the automation.

As stated in the previous research, Miller & Parasuraman [25] understand that a "mixture of human and automation involvement is frequently desirable, as compared with the extremes of full or no automation. Thus, human-machine systems must be designed for an appropriate relationship, allowing both parties to share responsibility, authority, and autonomy in a safe, efficient, and reliable fashion" [25]. How this should be achieved is however not sufficiently addressed in previous literature. This article is addressing this issue, with a focus on safety.

One of the contribution of this paper is the presented case study of stakeholders views on intelligent human–robot collaboration. It has demonstrated that industry stakeholders within the automotive industry are largely reluctant towards human–robot collaboration at the production facilities. They generally have bad experience of previous installations of cobots. However, they refer to applications where one small robot has been installed in a production line that is mainly manual. We argue that they do not evaluate human–robot collaboration as such but rather the functionality of an individual small robot that can be used without fences. But with weaker performance in the sense of lifting capacity, speed, and efficiency.

One of the major obstacles for introducing human–robot collaboration is that current regulation is understood by the stakeholders to make the collaborative automation installations slow and expensive. The process of making CE-markings is also understood as complicated, especially if robots are to be used in a flexible way. In our understanding the current regulation and safety standards are not well adapted to robots operating in the same physical space as humans. This is due to the fact that they are based on a traditional view on robotic automation.

To conclude, several attempts have been made to facilitate safe use of human–robot collaborative systems. However, we argue that i) existing categories and conceptualizations used to guide safety evaluation are insufficient, because they do not provide an all encompassing taxonomy relevant to challenges of facilitating intelligent human robot collaboration in industrial settings ii) the intelligent and autonomous aspect of intelligent human robot collaboration is not sufficiently addressed in existing regulation and standards, making the regulation inadequate for efficient and flexible solutions iii) current standards do not enable evaluation of the trade off between safety, efficiency and flexibility, and (iv) the regulation has a lack of focus on active safety and using the control system as a safety measure.

Regarding current safety standards, they can be used with simple cobot installations but the classification of collaborative operation modes and safety evaluation are much adapted to traditional automation solution They do not allow flexibility nor considering performance requirements, since they are mainly eliminating risks by limiting speed of force, or enforcing distance. These limitations can be an obstacle for solutions that involve real collaboration, which require higher level of intelligence and autonomy.

As a tentative solution to some of the concluded problems above, we proposed the new approach *deliberative safety* and contribute to classification of new safety measures to enable safety evaluation based on a number of fundamental KPIs: degree of freedom, interaction, speed, concurrency, coordination, monitoring and complexity. These KPIs can determine the level of flexibility, the level of efficiency, the level of interaction, how and when to share a space and the level of complexity required for implementing each of the suggested safety measures.

While the first measures labeled as "perimeter safety", "zone safety" and "reactive safety" fit well together with current risk evaluation in the robotic standards based on the Machinery Directive for simpler operation, the "planned safety" and "active safety" measures are more challenging to apply, advocating more reasoning between human and robots and shared decision making in truly collaborative operations. Using the deliberative safety approach and allowing adjustable safety

measures in operation will enable new possibilities. However, this will requires new types of intelligent and adaptive control systems, a reliable communication architecture, advanced monitoring, and an intuitive human-machine interface.

Inevitable, using the deliberative safety concept may incorporate new type of risks. Therefore it is necessary to develop new risk assessment and risk reduction methods that analyze and evaluate risks due to intelligent software and hardware to ensure that these new safety measures are able to handle or reduce the upcoming risks. The deliberative safety concept might require more knowledgeable human operators and new type of training to deal with new ways of thinking about risks an responsibility of the system.

The deliberative safety concept will also require changes in current regulation and safety standards, as well as in the way safety is dealt with by government agencies. The development of technology might challenge existing paradigms of regulation and division of responsibility in society; this is true for several AI-driven applications used within different fields of society. Specifically, the introduction of intelligent system (including AI) in organizations that we are dealing with in this paper might challenge established ways of organizing responsibility in decision-making, work management responsibility, and labor law responsibility linked to non-human entities acting autonomously. The development of suitable regulation is a topic for further inquiry in the scholarly literature as well as among government agencies and other involved stakeholder organizations. However, creating and developing standardization of technologies is a long term effort and therefore we aim together with the involved industry in this paper to assess the feasibility and the value of the suggested safety measures, first on several use cases in a lab environment and second on real cases in the production, to start the process of updating current standards to also incorporate these new intelligent human–robot collaboration system.

To be able to update the current industrial standards, the following research questions must be answered in the future: *How can human and robots express and understand each others intentions to be able to deliberate in planning for the best safety measures and in acting to support each others to avoid and handle risks?* and, *How can current risk assessment procedure be adapted to include deliberation in intelligent systems and to evaluate the proposed safety measures?* This will be addressed by investigating different human interaction strategies using the latest advance in User Experience Design to enable the deliberative concept and define how to handle changes and uncertainties but also to adapt and develop risk assessment and risk reduction methods that encompasses risks due to intelligence and autonomy.

Finally, a concluding remark is that many involved stakeholders are strongly influenced by how traditional automation is functioning when evaluating human–robot collaborative solutions, both in concern to evaluating safety and efficiency. An argument in this article is that it will require a radical change in perception and mentality as well as in action to be able to utilize the full potential of intelligent human–robot interaction enabling flexibility and efficiency.

CRediT authorship contribution statement

Atieh Hanna: Drafted, Revised, Reviewed, Approved the final version. **Simon Larsson:** Drafted, Revised, Reviewed, Approved the final version. **Per-Lage Götvall:** Drafted, Revised, Reviewed, Approved the final version. **Kristofer Bengtsson:** Drafted, Revised, Reviewed, Approved the final version.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Valeria Villani, Fabio Pini, Francesco Leali, Cristian Secchi, Survey on human-robot collaboration in industrial settings: Safety, intuitive interfaces and applications, 2018, URL http://dx.doi.org/10.1016/j.mechatronics.2018.02.009.

[2] Ericsson, Creative machines: How artificial intelligence will impact the future labor market, 2021, https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/creative-machines. (Accessed 02 June 2021).

[3] Ahmed R. Sadik, Bodo Urban, CPROSA-holarchy: An enhanced PROSA model to enable worker–cobot agile manufacturing, Int. J. Mech. Eng. Robot. Res. 7 (3) (2018) 296–304.

[4] ISO/TS 15066:2016 – robots and robotic devices – collaborative robots, 2016, https://www.iso.org/obp/ui/#iso:std:iso:ts:15066:ed-1:v1:en.

[5] Atieh Hanna, Towards Intelligent and Collaborative Automation of Automotive Final Assembly, Mälardalen University, 2021.

[6] J. Habermas, Ciaran P. Cronin, P.D. Greiff, The inclusion of the other: Studies in political theory, 1999.

[7] Wikipedia, Deliberation, 2021, https://en.wikipedia.org/wiki/Deliberation. (Accessed 19 June 2021).

[8] Ken Goldberg, Robots and the return to collaborative intelligence, Nat. Mach. Intell. 1 (1) (2019) 2–4, http://dx.doi.org/10.1038/s42256-018-0008-x.

[9] T.W. Malone, Superminds: The Surprising Power of People and Computers Thinking Together, Little, Brown, 2018, URL https://books.google.se/books?id=Qe0zDwAAQBAJ.

[10] Erik Hollnagel, Robert L. Wears, Jeffrey Braithwaite, From Safety-I to Safety-II: A White Paper, Technical Report, Odense: University of Southern Denmark; Gainesville: University of Florida; Macquarie Park, NSW: Macquarie University, 2015.

[11] B. Alenljung, J. Lindblom, User Experience of Socially Interactive Robots: Its Role and Relevance, Handbook of Research on Synthesizing Human Emotion in Intelligent Systems and Robotics, edited by Jordi Vallverdú. IGIGlobal, 2015.

[12] Jessica Lindblom, Beatrice Alenljung, Erik Billing, Evaluating the user experience of human–robot interaction, in: Céline Jost, Brigitte Le Pévédic, Tony Belpaeme, Cindy Bethel, Dimitrios Chrysostomou, Nigel Crook, Marine Grandgeorge, Nicole Mirnig (Eds.), Human-Robot Interaction: Evaluation Methods and their Standardization, Springer International Publishing, Cham, 2020, pp. 231–256, http://dx.doi.org/10.1007/978-3-030-42307-0_9.

[13] Lihui Wang, R Gao, József Váncza, J Krüger, Xi Vincent Wang, S Makris, G Chryssolouris, Symbiotic human-robot collaborative assembly, CIRP Ann. 68 (2) (2019) 701–726.

[14] Jörg Krüger, Terje K. Lien, Alexander Verl, Cooperation of human and machines in assembly lines, CIRP Ann. 58 (2) (2009) 628–646.

[15] Carole S Franklin, Elena G Dominguez, Jeff D Fryman, Mark L Lewandowski, Collaborative robotics: New era of human-robot cooperation in the workplace q,

[16] José Saenz, Norbert Elkmann, Olivier Gibaru, Pedro Neto, Survey of methods for design of collaborative robotics applications- why safety is a barrier to more widespread robotics uptake, in: ACM International Conference Proceeding Series. Part F137690, Association for Computing Machinery, 2018, pp. 95–101, http://dx.doi.org/10.1145/3191477.3191507.

[17] Przemyslaw A. Lasota, Terrence Fong, A Survey of Methods for Safe Human-Robot Interaction. Vol. 5, 2017, pp. 261–349, http://dx.doi.org/10.1561/2300000052.

[18] Luca Gualtieri, Erwin Rauch, Renato Vidoni, Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review, Robot. Comput.-Integr. Manuf. 67 (2021) http://dx.doi.org/10.1016/j.rcim.2020.101998.

[19] Federico Vicentini, Terminology in safety of collaborative robotics, 2020, http://dx.doi.org/10.1016/j.rcim.2019.101921.

[20] Guide to application of the Machinery Directive 2006/42/EC - Edition 2.2, URL https://ec.europa.eu/.

[21] SS-ISO 10218-1:2011 – robots and robotic devices – safety requirements for industrial robots – Part 1: Robots, 2011, URL https://www.iso.org/obp/ui/#iso:std:iso:10218:-1:ed-2:v1:en.

[22] SS-ISO 10218-2:2011 – robots and robotic devices – safety requirements for industrial robots – Part 2: Robot system and integration, 2011, URL https://www.iso.org/obp/ui/#iso:std:iso:10218:-2:ed-1:v1:en.

[23] Luca Gualtieri, Erwin Rauch, Renato Vidoni, Dominik T Matt, Safety, ergonomics and efficiency in human-robot collaborative assembly: design guidelines and requirements, Procedia CIRP 91 (2020) 367–372.

[24] Riccardo Gervasi, Luca Mastrogiacomo, Fiorenzo Franceschini, A conceptual framework to evaluate human-robot collaboration, Int. J. Adv. Manuf. Technol. 108 (2020) 841–865.

[25] Christopher A. Miller, Raja Parasuraman, Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control, Hum. Factors 49 (1) (2007) 57–75.

[26] Deneen M. Hatmaker, Engineering identity: Gender and professional identity negotiation among women engineers, Gender, Work Org. 20 (4) (2013) 382–396.

[27] Simon Larsson, Magnus Jansson, Åsa Boholm, Expert stakeholders' perception of nanotechnology: risk, benefit, knowledge, and regulation, J. Nanoparticle Res. 21 (3) (2019) 1–17.

[28] Simon Larsson, Kristofer Bengtsson, Enabling Human-Robot Collaboration and Intelligent Automation in the Automotive Industry: A Study of Stakeholder Perspectives, Technical Report, Gothenburg Research Institute, GRI report 2022:1, 2022, URL https://gupea.ub.gu.se/handle/2077/71636.

[29] Steinar Kvale, Svend Brinkmann, Interviews: Learning the Craft of Qualitative Research Interviewing, sage, 2009.

[30] Barbara DiCicco-Bloom, Benjamin F. Crabtree, The qualitative research interview, Med. Educ. 40 (4) (2006) 314–321.

[31] Juliet M. Corbin, Anselm Strauss, Grounded theory research: Procedures, canons, and evaluative criteria, Qual. Sociol. 13 (1) (1990) 3–21.

[32] Gery W. Ryan, H. Russell Bernard, Techniques to identify themes, Field Methods 15 (1) (2003) 85–109.

[33] Toshiyuki Inagaki, et al., Adaptive automation: Sharing and trading of control, Handb. Cogn. Task Des. 8 (2003) 147–169.

[34] William B. Rouse, Adaptive aiding for human/computer control, Hum. Factors 30 (4) (1988) 431–443.

[35] Mark W. Scerbo, Theoretical perspectives on adaptive automation, in: Automation and Human Performance: Theory and Applications, CRC Press, 2018, pp. 37–63.

[36] Terrin Babu Pulikottil, Stefania Pellegrinelli, Nicola Pedrocchi, A software tool for human-robot shared-workspace collaboration with task precedence constraints, Robot. Comput.-Integr. Manuf. 67 (2021) 102051, http://dx.doi.org/10.1016/j.rcim.2020.102051, URL https://www.sciencedirect.com/science/article/pii/S0736584520302623.

[37] George Michalos, Niki Kousi, Panagiotis Karagiannis, Christos Gkournelos, Konstantinos Dimoulas, Spyridon Koukas, Konstantinos Mparis, Apostolis Papavasileiou, Sotiris Makris, Seamless human robot collaborative assembly – an automotive case study, Mechatronics 55 (2018) 194–211, http://dx.doi.org/10.1016/j.mechatronics.2018.08.006, URL https://www.sciencedirect.com/science/article/pii/S0957415818301326.

[38] Emmanuel Dean-Leon, Karinne Ramirez-Amaro, Florian Bergner, Ilya Dianov, Gordon Cheng, Integration of robotic technologies for rapidly deployable robots, IEEE Trans. Ind. Inf. 14 (4) (2018) 1691–1700, http://dx.doi.org/10.1109/TII.2017.2766096.

[39] Shozo Takata, Takeo Hirano, Human and robot allocation method for hybrid assembly systems, CIRP Ann. 60 (1) (2011) 9–12.

[40] Morteza Dianatfar, Jyrki Latokartano, Minna Lanz, Task balancing between human and robot in mid-heavy assembly tasks, Procedia CIRP 81 (2019) 157–161.