ICCEPM 2024 https://dx.doi.org/10.6106/ICCEPM.2024.1088
The 10th International Conference on Construction Engineering and Project Management
Jul. 29-Aug.1, 2024, Sapporo

Errors in Human-Robot Interaction Accidents: A Taxonomy and Network Analysis

Brian H.W. GUO[1]*, Yonger ZUO[2], Yang Miang Goh[3], Jae-Yong Lim[4]

[1] *Department of Civil and Natural Resources Engineering, Faculty of Engineering, University of Canterbury, New Zealand,* E-mail address: brian.guo@canterbury.ac.nz
[2] *Department of Civil and Natural Resources Engineering, Faculty of Engineering, University of Canterbury, New Zealand,* E-mail address: yonger.zuo@pg.canterbury.ac.nz
[3] *Department of Built Environment, National University of Singapore, Singapore,* E-mail address: bdggym@nus.edu.sg
[4] *Department of Safety Engineering, Seoul National University of Science and Technology, 232, Gongneung-ro, Nowon-gu, Seoul, 01811, Republic of Korea, ,* E-mail address: jylim@seoultech.ac.kr

Abstract:

To enhance the investigation, analysis, and design of safe human-robot interactions (HRI), this study develops a comprehensive taxonomy of safety-related errors in HRI and examines the relationships between errors and the types and levels of HRI. Analyzing 262 HRI accident case reports, the research identifies and categorizes human and robot errors through qualitative analysis. The resulting taxonomy divides human errors into procedure, intrusion, operation, and situation awareness errors, and robot errors into system and safeguarding failures, operational errors, and design flaws. A network of human and robot errors was developed by applying Gephi to represent the human-robot error interactions. The results indicated that "misjudgment of the robot's operational status," "inadvertent activation of the robot," "working within an energized robotic cell without adequate safety measures," and "failure to deenergize/stop the robot" are among those most frequently linked to robot errors. "Inadequate lockout/tagout" and "absence of human detection and protective stop functions" stand out as the most frequent human-robot error interaction.

Key words: human-robot interaction, human error, robotics, workplace safety

1. INTRODUCTION

Human error has been one of the central topics in safety science during the past five decades. In the construction safety research domain, early research focused on modeling and examining unsafe behavior by drawing upon theories developed in organization science and behavior science [1,2]. Inspired by Reason's Swiss cheese model [3] and new constructs derived from organizational science, such as safety culture [4] and safety climate [5], researchers made efforts to examine the nature of human error on construction sites [2,6–8]. Research advances made in the research stream significantly improved the understanding of how organizational factors impact safety performance at the group and individual levels. Behavioral-based safety (BBS) techniques were also applied and examined in the construction context [9–11]. Several behavior and motivation theories (e.g., goal-setting theory and theory of planned behavior) were applied to interpret why workers behave unsafely on construction sites.

In the rapidly evolving landscape of industrial technology, human-robot interaction (HRI) has emerged as a critical area of research and development. As industries increasingly deploy robotic systems to enhance efficiency and productivity, understanding the dynamics of human-robot collaboration becomes essential. The advent of advanced robotics in industries marked a paradigm shift in how tasks are performed, ranging from manufacturing to service delivery. Robots, equipped with artificial intelligence and machine learning capabilities, are no longer confined to repetitive, menial tasks. Instead, they are now capable of performing complex functions that require adaptive and intelligent behavior, often working alongside human counterparts.

However, this integration of human and robotic agents in a shared workspace is not without its challenges. The complexity of human-robot interaction introduces a myriad of potential errors, which have resulted in accidents and injuries. These errors can range from minor inefficiencies to significant safety hazards, thereby making their study a matter of paramount importance. Researchers have made efforts to develop taxonomies to categorize various dimensions of human-robot interactions. For example, [12] developed a HRI taxonomy that categorizes HRI in terms of interaction context, robot and team classification. [13] proposed a taxonomy for HRI which focuses on team composition and space-time location. In the construction industry, [14] proposed five categories were proposed in the taxonomy that characterize the interplay between robot autonomy and human effort, namely preprogramming, adaptive manipulation, imitation learning, improvisatory control, and full autonomy.

Despite their popularity, arguably, these taxonomies are too generic to be useful for safe human-robot interaction management and robotics design. There is an urgent need for a specific taxonomy of human-robot interaction errors. The absence of a comprehensive taxonomy of errors in human-robot interactions (HRI) represents a significant gap in the current body of research. This gap is not merely a missing piece in academic literature, but it has profound implications for the development, implementation, and optimization of robotic systems in industrial settings.

To fill the knowledge gap, this paper aims to (1) develop a taxonomy of human and robot errors in the HRI context and (2) investigate the relationships between errors and human-robot interactions combination types and levels.

2. LITERATURE REVIEW

Risk assessment of industrial robots is usually carried out according to the international standards ISO 12100 and ISO 14121-2. Existing system analysis methods applied in human-robot collaboration scenarios are designed to support hazard identification and risk assessment in standard procedures [15]. Established security engineering methods are widely used in human-robot collaboration scenarios. The safety engineering methods used in most of these studies are Failure Modes and Effects/Criticality Analysis (FMEA/FMECA), Fault Tree Analysis (FTA), and Hazard and Operability Studies (HAZOP). Since these methods were originally developed for traditional industries, which are often static or predictable, this is very different from the HRC environment. Therefore, to better adapt to human-robot collaboration scenarios, there are also some studies that combine these methods with UML diagrams, Early Warning Sign Analysis, etc. However, there are still limitations in the applicability of robots. There are also many subsequent studies to develop adaptive risk assessment tools specifically for human-robot collaboration scenarios and the field of robotics. Some researchers try to conduct human-robot collaboration system risk assessment from the perspectives of task decomposition, formal verification, and expert systems. But when it comes time to conduct an in-depth analysis of a complex scene, these angles don't seem to deliver the level of detail and accuracy required. Simulation-based security testing becomes an alternative risk assessment option. However, simulation tools rely on real-world data and many calculation processes, which may be difficult to cover all possible interaction scenarios and conditions, and difficult to provide real-time risk assessment.

Due to the vast number of human-robot systems, the potential for errors is also enormous, especially when considering the diversity of human-robot interactions. Therefore, identifying human and robotic errors in a system requires a taxonomy. The most common human error classification [16] and robot classification [17], etc., have been developed for other industries as well as traditional industrial robots. Although some research in the past has also used it for error analysis of human-robot collaboration, it is far from enough. In human-robot collaboration, initially, [18] proposed a classification method for robot physical errors and human errors based on the collaboration of robots such as unmanned ground vehicles (UGV). While this taxonomy is broad, there are other interaction failures that are not considered. For example, it does not account for other types of human error. Later, [19] proposed a human-robot error classification method divided into robotic technology errors and interaction errors. Interaction errors include human errors. Although this taxonomy establishes a detailed classification of robot errors from a technical perspective, human errors are still classified according to the established framework. They are not fully applicable to human-robot collaboration systems. Notably, the first classification of the causes of errors in human-robot systems was developed by [20]. This new classification model considers errors not only in the system's individual components but also arising from the interactions between them. However, this classification still lacks a more specific and objective quantification process.

3. METHOD

A total of 262 HRI accident case reports were collected from different sources, including 200 cases used in [21], 54 cases from the Occupational Safety and Health Administration (OSHA) database in the United States (US), and 8 cases from Google search. Each report provides details of the chain of events and causes of accidents. Each accident report was subjected to an initial review to ascertain its relevance and to extract preliminary data. This step ensured that the reports were directly related to HRI and contained sufficient detail for further analysis. Reports that met these criteria were then categorized based on the type of human-robot interaction, the industry context, and the nature of the accident.

To identify both human and robot errors, the following definitions were adopted. Human error refers to active and unsafe acts performed by frontline workers. Identification and analysis of human error in the HRI context was conceptually based on Rasmussen's SRK (Skill, Rule, Knowledge) model of human error [22]. It is beyond the scope of this paper to debate the "'old view' and 'new view' of human error" [23]. On the other hand, robot error refers to mistakes, malfunctions, and failures originating from the robot itself in terms of its hardware, software, and integrated machines and systems. It involves robot functions and behaviors that deviate from ISO standards. For example, ISO standard [24] requires that "*Every robot shall have a protective stop function and an independent emergency stop function*." Thus, "lack of a protective stop function" was coded as a robot error in relevant accident cases. In human-robot collaboration scenarios, both [24] and [25] require that "*The robot shall stop when a human is in the collaborative workspace*."

Each report was analyzed in depth to identify and code the specific human and robot errors involved. This analysis involved a line-by-line examination of the text to identify error instances, which were then coded according to the previously developed scheme. The coding process was iterative, allowing for the refinement of the coding scheme as new types of errors were identified. Based on the coded data, we constructed a taxonomy of human and robot errors in HRI. This taxonomy was structured to reflect the hierarchical and relational nature of the errors, from broad categories to specific subtypes. The development of the taxonomy was an iterative process involving continuous refinement and validation against the coded accident reports to ensure comprehensive coverage and logical structure.

To represent and visualize human-robot interaction errors, we employed Gephi, an open-source network analysis and visualization software tool [26], to construct and analyze a network representation of human and robot errors. Utilizing Gephi, we constructed a network where nodes represented individual errors, categorized into 'human' or 'robot' errors based on the originator. Edges between nodes were established based on the relationships identified during data preparation, such as a human error co-exists with a subsequent robot error in an accident. The relationship represents the interaction between human-robot errors. Edge weights were assigned to represent the frequency or severity of the error connections. We leveraged Gephi's visualization capabilities to create a graphical representation of the network. This included using node size and color coding to represent error severity and type, respectively, and adjusting edge thickness to reflect connection strength.

4. RESULTS

4.1. A taxonomy of human error

Human errors in the HRI context were categorized into four categories: procedure error, intrusion error, operation error, and situation awareness error, as shown in Figure 1. Each category is further divided into sub-categories to encapsulate the multifaceted nature of human errors in HRI. Below is an elaboration of each category and its sub-categories.

Procedure error: Procedure errors are violations or deviations from established protocols and safety measures during interactions with robotic systems. This includes the failure to implement adequate lockout/tagout practices, which ensure machinery is properly shut off and not started up unexpectedly during maintenance. It also covers inadequate entry and exit procedures for robotic cells, highlighting the importance of following specific protocols to ensure safety. The omission of de-energizing robots during maintenance tasks poses significant risks, as does working within an energized robotic cell without appropriate safety measures. Additionally, the failure to wear personal protective equipment (PPE) and errors in communication, such as starting a robot without informing nearby workers, fall under this category, emphasizing the need for stringent adherence to safety protocols.

Intrusion Error involves unauthorized or unintended access to robotic areas. Unauthorized access refers to individuals entering robotic workspaces without permission, potentially leading to disruptions or injuries. Similarly, intrusion through a light curtain, a safety mechanism designed to stop robot operation when breached, underscores the dangers of bypassing safety systems.

Operation Error focuses on mistakes made during the direct interaction with robots, such as errors with pendant controls, which are handheld devices used to manually operate robots. Incorrect inputs can lead to unintended robot movements. Errors during programming sessions can result in the robot performing unintended actions, and accidentally activating the robot can lead to unforeseen operations, highlighting the need for careful handling and clear protocols.

Situation Awareness Error refers to errors related to inattention and a lack of understanding to the robot's operational status and environment. Misjudging the robot's operational status, such as assuming it is inactive when it is not, can lead to unsafe interactions. A lack of knowledge about the programmed tasks or expected motions of the robot can result in operational inefficiencies or dangers. Misinterpreting the robot's operational behavior can increase the risk of accidents, and distracted attention from the robot's actions and surroundings can lead to critical oversights.

The taxonomy of human error in human-robot interaction aligns closely with Rasmussen's SRK (Skill, Rule, Knowledge) model of human error [22], which categorizes errors based on the level of cognitive processing involved. The linkages are presented in Table 1.

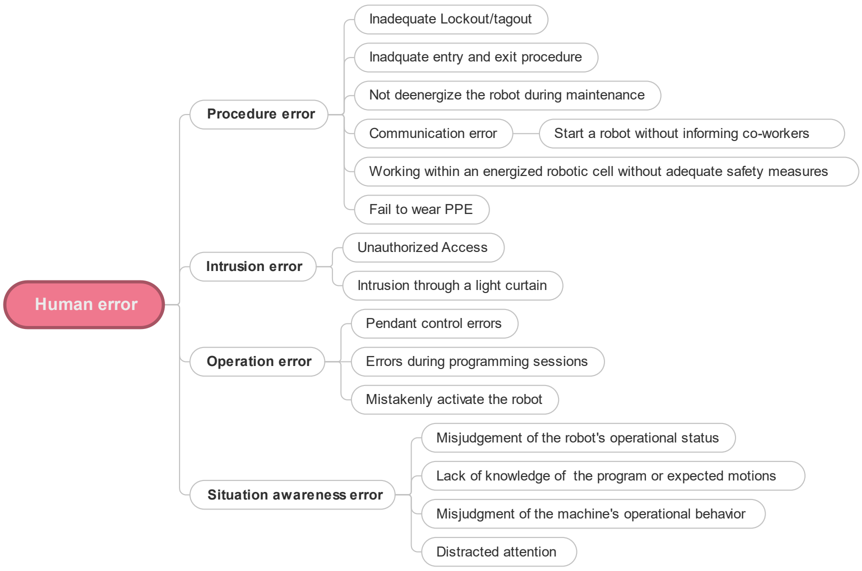

Figure 1 A taxonomy of human error in HRI

Table 1. SRK level of human error in HRI

SRK Level	Error category	Subcategories
Skill-based (S)	Operation error	Pendant Control Errors, Errors during Programming Sessions, Mistakenly Activate the Robot
	Situation awareness error	Distracted Attention
Rule-based (R)	Procedure Error	Inadequate Lockout/Tagout, Inadequate Entry and Exit Procedure, Not Deenergize the Robot During Maintenance, Working within an Energized Robotic Cell without Adequate Safety Measures, Fail to Wear PPE, Communication error
Knowledge-based (K)	Situation Awareness Error	Misjudgment of the Robot's Operational Status, Lack of Knowledge of the Program or Expected Motions, Misjudgment of the Machine's Operational Behavior

4.2. A taxonomy of robot error

Based on the analysis of the accident cases, robot errors are categorized into four main types: system failures, safeguarding failures, operational errors, and design flaws (see Figure 2). Details of each type are presented as follows.

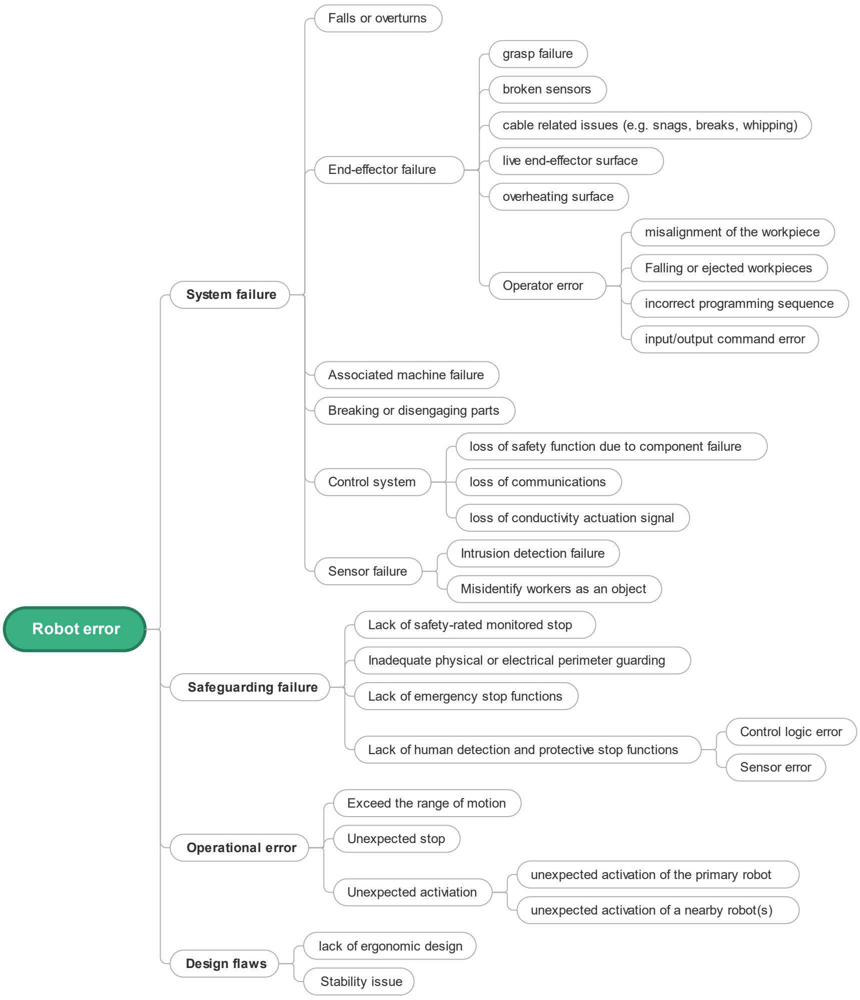

Figure 2 A taxonomy of robot error in HRI

System failures refer to errors arising from the malfunctioning of the robot's core components. This includes failures within the control system, such as the loss of safety functions due to component failure, communication breakdowns, and issues with actuation signal conductivity. Associated machine failures, part disengagements, robot falls or overturns, sensor malfunctions, and end-effector failures also fall under this category. Note that sensor failures can lead to the inability to detect intrusions properly or the misidentification of workers as objects. End-effector failures encompass a range of issues, from grasp failures to overheating surfaces, often exacerbated by operator errors like misalignment of workpieces or incorrect programming sequences.

Safeguarding Failures highlight deficiencies in the protective measures designed to ensure human safety around robots. This encompasses the lack of safety-rated monitored stops, inadequate guarding both physically and electrically, and the absence of crucial emergency or protective stop functionalities. Specifically, the lack of emergency stop functions can stem from control logic errors or sensor inaccuracies, posing significant risks to human operators.

Operational errors involve the robot's performance deviating from expected or programmed actions. This includes exceeding the robot's designed range of motion, unexpected activations, and sudden, unanticipated stops. Unexpected activations are particularly concerning, whether it involves the primary robot or nearby robotic systems, as they can lead to uncontrolled movements and potential safety hazards.

Design flaws address errors stemming from the initial design and engineering of the robot. This includes the lack of ergonomic considerations in the robot's design, which can affect human interaction and efficiency, and stability issues that may lead to robot falls or compromised performance.

4.3. A network of human and robot errors

Figure 3 presents a network visualization of human and robot errors, including only those nodes with a degree greater than two for the sake of succinctness. The size of each node reflects its degree within the network, which indicates the quantity of connections it has with other nodes. Note that errors such as "misjudgment of the robot's operational status," "inadvertent activation of the robot," "working within an energized robotic cell without adequate safety measures," and "failure to de-energize/stop the robot" are among those most frequently linked to robot errors. For instance, the error "misjudgment of the robot's operational status" was associated with nine different robot errors in the accidents analyzed. On the other hand, "unexpected activation/movement of the primary robot," "absence of human detection and protective stop functions," and "inadequate physical or electrical perimeter guarding" emerged as the most prevalent and critical robot errors within the network.

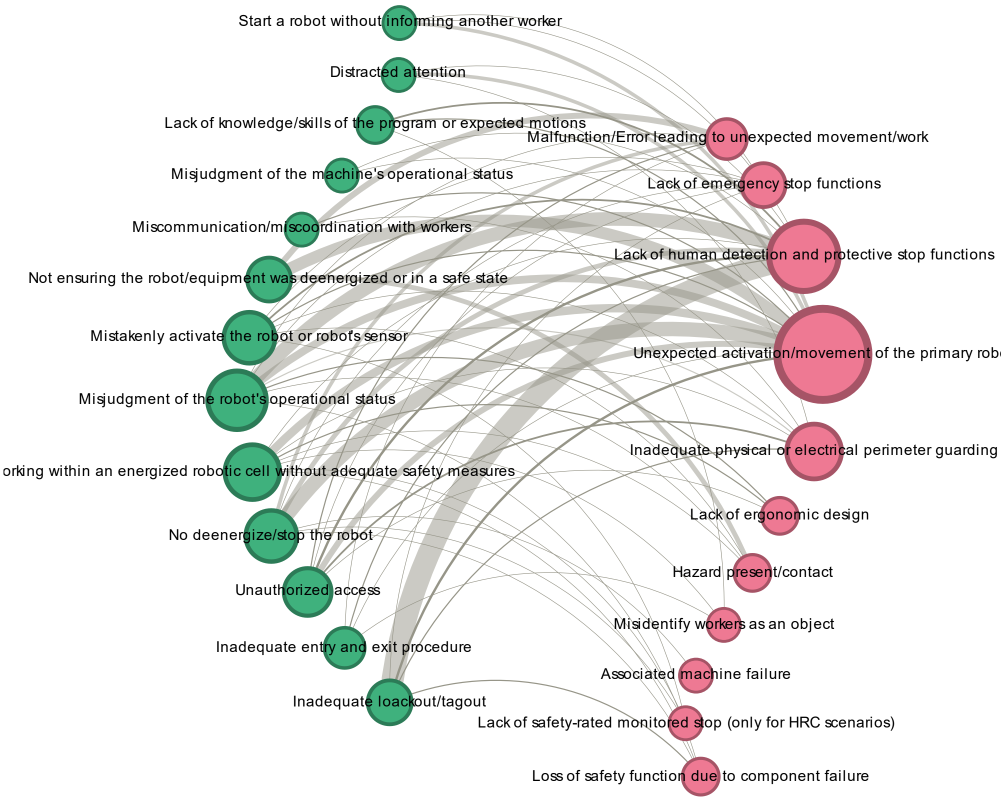

Figure 3 A network of human and robot errors

Connections between nodes, representing the simultaneous contribution of human and robot errors to accidents, are visualized through edges. The thickness of these edges represents the frequency and intensity of their co-occurrence across all examined accident cases. The connection between "inadequate lockout/tagout" and "absence of human detection and protective stop functions" stands out as particularly common, illustrating a typical scenario in human-robot interactions where an accident occurs because a worker enters a robotic work cell without adhering to lockout/tagout protocols, and the robot fails to cease operations due to its inability to detect the worker's presence. This scenario is often compounded by another human error, "misjudgment of the robot's operational status," especially in instances where the robot appears inactive prior to the worker's entry into the cell.

5. DISCUSSION AND CONCLUSION

This paper aimed to develop a taxonomy of human and robot errors in the HRI context. The taxonomy of human and robot errors developed in this study significantly contributes to the field of safety science, particularly within the context of HRI in industrial settings. By categorizing and analyzing the various types of errors that can occur during HRI, this research provides a structured framework that can be instrumental in understanding the complex dynamics between humans and robots working in close proximity. The findings of this study underscore the critical interplay between human errors and robot malfunctions, highlighting scenarios where the lack of adequate safeguards and error-tolerant design can lead to accidents. For instance, the frequent co-existence of "inadequate lockout/tagout" procedures with "lack of human detection and protective stop functions" in robots points to a systemic vulnerability in current HRI setups. This aligns with the Joint Cognitive Systems (JCS) principle of "Defense in Depth," [23] suggesting that multiple layers of defense, including procedural safeguards and technological fail-safes, are essential to prevent the escalation of such incidents. Moreover, the network analysis of human and robot errors, facilitated by Gephi visualization, reveals critical nodes where interventions could significantly enhance safety. Addressing errors like "misjudgment of the robot's operational status" and "working within an energized robotic cell without adequate safety measures" through better training, clearer protocols, and enhanced situational awareness tools can help mitigate the risks associated with human errors.

This paper has the following limitations. First, as the accident reports do not consist of latent failures related to the supervision and management level, this has restricted our analysis to active human and robot errors that occur at the operational level. However, it is important to clarify that this paper does not suggest that human and robot errors are the sole focus of accident analysis and investigation. Future research should aim to contextualize these errors comprehensively and trace their causal links to underlying factors. Second, the details available in the accident reports are insufficient for examining the reasons behind the workers' errors from the perspectives of cognitive psychology, behavior, and management. In addition, it should be noted that all robot errors can be attributed to technological design and managerial issues.

REFERENCES

[1] B.H. Guo, T.W. Yiu, V.A. González, Identifying behaviour patterns of construction safety using system archetypes, Accident Analysis & Prevention 80 (2015) 125–141.

[2] Y.M. Goh, N.F. Binte Sa'adon, Cognitive factors influencing safety behavior at height: a multimethod exploratory study, Journal of Construction Engineering and Management 141 (2015) 04015003.

[3] J. Reason, Managing the risks of organizational accidents, Routledge, 2016.

[4] F.W. Guldenmund, The nature of safety culture: a review of theory and research, Safety Science 34 (2000) 215–257.

[5] D. Zohar, Safety climate in industrial organizations: theoretical and applied implications., Journal of Applied Psychology 65 (1980) 96.

[6] G.J. Fogarty, A. Shaw, Safety climate and the theory of planned behavior: Towards the prediction of unsafe behavior, Accident Analysis & Prevention 42 (2010) 1455–1459.

[7] B.H. Guo, T.W. Yiu, V.A. González, Predicting safety behavior in the construction industry: Development and test of an integrative model, Safety Science 84 (2016) 1–11.

[8] S. Mohamed, Safety climate in construction site environments, Journal of Construction Engineering and Management 128 (2002) 375–384.
[9] R.M. Choudhry, Behavior-based safety on construction sites: A case study, Accident Analysis & Prevention 70 (2014) 14–23.
[10] A.R. Duff, I.T. Robertson, R.A. Phillips, M.D. Cooper, Improving safety by the modification of behaviour, Construction Management and Economics 12 (1994) 67–78.
[11] B.H. Guo, Y.M. Goh, K.L.X. Wong, A system dynamics view of a behavior-based safety program in the construction industry, Safety Science 104 (2018) 202–215.
[12] L. Onnasch, E. Roesler, A Taxonomy to Structure and Analyze Human–Robot Interaction, Int J of Soc Robotics 13 (2021) 833–849. https://doi.org/10.1007/s12369-020-00666-5.
[13] H.A. Yanco, J.L. Drury, A taxonomy for human-robot interaction, in: Proceedings of the AAAI Fall Symposium on Human-Robot Interaction, 2002: pp. 111–119. https://cdn.aaai.org/Symposia/Fall/2002/FS-02-03/FS02-03-015.pdf (accessed December 21, 2023).
[14] C.-J. Liang, X. Wang, V.R. Kamat, C.C. Menassa, Human–Robot Collaboration in Construction: Classification and Research Trends, J. Constr. Eng. Manage. 147 (2021) 03121006. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002154.
[15] T.P. Huck, N. Münch, L. Hornung, C. Ledermann, C. Wurll, Risk assessment tools for industrial human-robot collaboration: Novel approaches and practical needs, Safety Science 141 (2021) 105288. https://doi.org/10.1016/j.ssci.2021.105288.
[16] J. Reason, Human Error, Cambridge University Press, 1990.
[17] A. Nakamura, K. Nagata, K. Harada, N. Yamanobe, Estimation and Categorization of Errors in Error Recovery Using Task Stratification and Error Classification, Journal of Robotics, Networking and Artificial Life 4 (2017) 163. https://doi.org/10.2991/jrnal.2017.4.2.13.
[18] J. Carlson, R.R. Murphy, How UGVs physically fail in the field, IEEE Transactions on Robotics 21 (2005) 423–437.
[19] S. Honig, T. Oron-Gilad, Understanding and resolving failures in human-robot interaction: Literature review and model development, Frontiers in Psychology 9 (2018) 861.
[20] A. Žunjić, Classification of Causes of Errors in the Human-Robot System, Journal of Robotics and Mechanical Engineering Research 2 (2018) 7–13.
[21] K. Lee, J. Shin, J.-Y. Lim, Critical hazard factors in the risk assessments of industrial robots: causal analysis and case studies, Safety and Health at Work 12 (2021) 496–504.
[22] J. Rasmussen, Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models, IEEE Transactions on Systems, Man, and Cybernetics (1983) 257–266.
[23] J.C. Le Coze, The 'new view'of human error. Origins, ambiguities, successes and critiques, Safety Science 154 (2022) 105853.
[24] S. ISO, Robots and Robotic Devices-Safety Requirements for Industrial Robots-Part 1: Robots (ISO 10218-1: 2012), International Organization for Standardization (2012).
[25] S. ISO, Robots and Robotic Devices—Collaborative Robots (ISO-15066: 2016), International Organization for Standardization (2016).
[26] M. Bastian, S. Heymann, M. Jacomy, Gephi: an open source software for exploring and manipulating networks, in: Proceedings of the International AAAI Conference on Web and Social Media, 2009: pp. 361–362. http://ojs.aaai.org/index.php/ICWSM/article/view/13937 (accessed February 7, 2024).