Human Factors in Design, Engineering, and Computing, Vol. 159, 2024, 793–804
https://doi.org/10.54941/ahfe1005646

Enhancing Emergency Response: Reliability Analysis of Human-Robot Collaboration

Daigo Misaki and Kazuma Shirakawa

Faculty of Engineering Department of Mechanical Systems Engineering, Kogakuin University, 1-21-6 Nishishinjuku Shinjuku-ku, Tokyo 1638677, Japan

ABSTRACT

The adoption of robots in daily life, such as service robots, is progressing and necessitates that humans make informed decisions when interacting with them. However, the relationships between humans and robots, particularly in emergencies, are not as developed as human-to-human relationships. A lack of understanding about robots often leads to significant accidents. To facilitate effective and appropriate collaboration, the analysis of human-robot interaction (HRI) is essential. This study focuses on analyzing "reliability," which is particularly crucial in the healthcare and training fields. We specifically examined the interactions between humans and robots during emergencies and analyzed the reliability of these interactions. Our verification method combines case studies and empirical experiments, beginning with a case analysis and followed by an experimental design based on these findings. Empirical experiments confirmed that combining visual and tactile feedback significantly affects interface reliability in HRIs. Designing empirical experiments based on case study results is a crucial analytical approach for enhancing the utility of services and healthcare robots for users.

Keywords: Human factor, Human robot collaboration, Reliability, Haptics,

INTRODUCTION

Designing new technologies and human-machine interfaces (HMIs) to prevent human errors is crucial for the success of future collaborative societies. Examples of the use of HMI, augmented reality (AR), robots, and artificial intelligence (AI) to enhance safety and efficiency can be found in various fields, including nuclear power generation (Anokhin et al., 2018), inspection drones (Riku Tsunori et al., 2019), healthcare (Dasho et al., 2022) (Nakamura et al., 2021) (Yamazaki et al., 2021), and aviation (Huseyin Avsar et al., 2016) (Njolomole et al., 2021). Designing reliable systems and robots is vital, particularly in healthcare, where research has focused on the reliability and dependability of AI of mobile robots (Asan et al., 2020) (Sahoo et al., 2023). Human errors often result from a multitude of factors, including communication between humans and machines, environmental conditions, and decision-making processes (Moghim et al., 2023). Therefore, understanding and designing human behavior and system interfaces are crucial for accident prevention (Moura et al., 2014). Moreover, individuals who commit human errors are often unaware of their mistakes and may require external cues to recognize their errors (Matsuo, 2009). Consequently, analyzing the factors contributing to human-machine communication errors in accident cases can significantly improve trust in collaborative environments and shed light on human decision-making characteristics.

Various studies have used text mining to analyze accident factors within single domains (Luo et al., 2021), explored human-AI collaboration potential (Lee et al., 2011), and investigated how cultural backgrounds affect ethical decision-making (Awad et al., 2018). However, research that systematically analyzes the causes of accidents from various fields and design perspectives remains limited. Additionally, there is insufficient discussion on the challenges arising in contemporary society owing to the increasing collaboration between humans and AI in collaborative environments.

To address these gaps, we conducted a case study to investigate accidents caused by human error in the aviation, railway, and marine domains in both Japan and Taiwan. This study considered the influence of cultural and societal backgrounds on accident causes, to reveal patterns and variations in accidents. We generated word clouds, as shown in Figure 1, and co-occurrence networks, as shown in Figure 2, to analyze the root causes. Our analysis of Japanese data (JTSB) [Japan Transport Safety Board (https://www.mlit.go.jp/jtsb/english.html) [Accessed February 5, 2024]] indicates a higher frequency of communication errors in the aviation sector, accidents due to time pressure in the railway sector, and errors resulting from overconfidence in the marine sector, which we attribute to variations in the adoption of automation technologies. In the aviation sector, where automation has advanced, human communication errors are prevalent because evaluations are often delegated to machines. By contrast, the marine sector, where automation is less prevalent, sees more errors related to individual biases and overconfidence. From Taiwanese data (TTSB) [Taiwan Transportation Safety Board (https://www.ttsb.gov.tw/english/) [Accessed February 5, 2024]], we observed that language differences were a prevalent factor contributing to errors, suggesting cultural differences compared with Japanese accident cases (Table 1). In our case study, human errors in accidents were notably influenced by individual backgrounds and experiences. Therefore, to reduce human errors in real-world coworking environments, we propose the development of appropriate information displays and collaborative behavior among humans, vehicles, and robots. Chang et al. (2022) researched enhancing communication between humans and vehicles using eye-tracking technology with the direct aim of reducing traffic accident risks in autonomous vehicles. Moreover, Miller (2016) introduced “Trust Fall” to determine if individuals trust autonomous systems in safe scenarios or rely on their judgment in crises. Verberne et al. (2015) define trust as the willingness to embrace vulnerabilities expecting positive outcomes from an entity or system.

Figure 1: Domains of human systems integration. (Adapted from U.S Air Force, 2005).

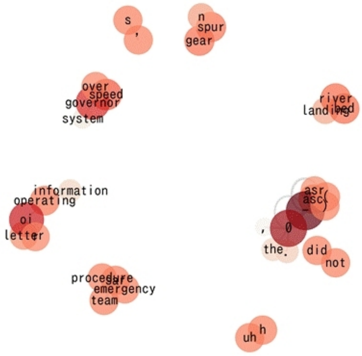

Figure 2: Co-occurrence network of accident cases in one TTSB data.

Table 1. Typical characteristics from the accident investigation reports by Japan Transport Safety Board (JTSB) and Taiwan Transportation Safety Board (TTSB).

	Japan (JTSB)	Taiwan (TTSB)
Aviation	Communication errors between people are high (6/16). Overreliances on experience errors are few (1/16).	Mistakes due to language differences in the manual.
Railway	Many mistakes related to time (7/8).	Mistakes due to differences in cultural backgrounds.
Marine	Many mistakes due to attention elsewhere or overreliance on experience (7/13).	Mistakes due to differences in cultural backgrounds.

Various fields, other than automobiles, employ numerous information display methods. In the medical field, haptic feedback from surgical robots is used for some applications, whereas other initiatives focus on robots that mimic human movements to enhance human-robot interactions (Takano et al., 2022) (Teppei Tsujita et al., 2022) (Hsieh et al., 2020). Discussions further highlight the critical role of tactile feedback in enhancing collaborative HRIs during operations (Jensen et al., 2021). Intuitive and user-friendly interfaces are highly preferred in environments where humans, machines, or systems collaborate, and intuitive and user-friendly interfaces are highly preferred (Miura et al., 2015). Therefore, incorporating human factors into the interface design is crucial for mitigating human error. This study concentrates on “mistakes in human assumptions,” evaluating human behavioral characteristics and information display elements that influence decision-making when actions deviate from intuition. We achieve this through a combination of case studies and empirical research. Furthermore, we propose specific strategies to minimize human errors in human-machine collaborative environments.

METHOD

This study focused on errors occurring during the interpretation and comparison phases, drawing on seven stages of action (Figure 3) (Norman, 1989). The evaluative phase consists of perception, interpretation, and comparison, which are phases prone to significant errors when one’s intuition or experience diverges from the external reality. Several evaluative factors exist between interpretation and comparison, often intuitively employing raw data as judgment elements as one gains experience with various situations. However, trust elements that evolve with experience lead to machine- or system-presented information being generally more accurate and having a lower error rate than human judgment.

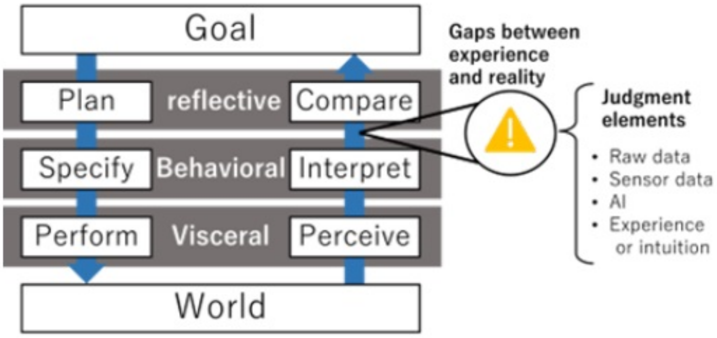

Figure 3: The seven stages of action.

Discrepancies between machine and system intentions and personal judgment may result in incorrect decisions, highlighting the importance of valuing data-driven decision-making. Furthermore, given the pace of technological advancement, quick adaptation to, and safe use of unfamiliar technology, Figure 4 depicts the experimental conditions that integrate these scenarios. Initially, the participants were divided into two groups: one experiencing anticipated movements and the other experiencing unexpected movements during robot operation. Subsequently, each group conducted operations first by sight alone, and then with the addition of sensor (vibration) feedback, totaling four robot operation trials. After every two trials, the participants were asked to complete a survey to gather feedback. We selected vibration as sensor feedback because of its direct and tangible impact on participants. Thus, by integrating these variables, we suggest experiments that account for scenarios that diverge from intuition or involve limited experience.

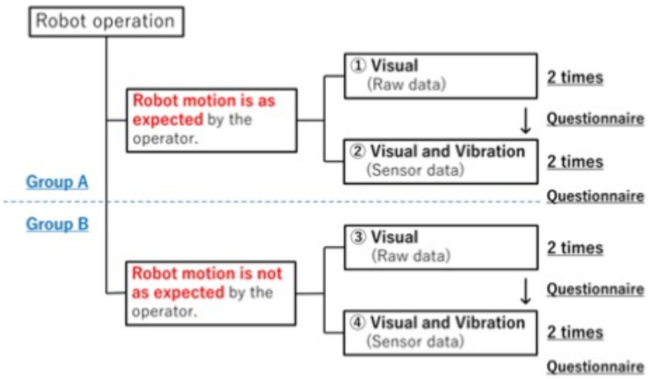

Figure 4: Experimental conditions.

As shown in Figure 5, the system first transmitted information from the analog stick of the game controller to the PC, enabling the participants to control the robot visually for conditions ① in Figure 4 and ③. For conditions ② and ④, which combine visual control with sensor (vibration) feedback, the controller was programmed to vibrate at specific positions of the analog stick to simulate tactile feedback. A vibration motor was strategically placed behind the controller to maximize the tactile feedback experienced by the user. Prototyping revealed that a single motor was sufficient to vibrate the entire controller effectively, leading to the decision to use only one motor. The operations relied solely on the analog stick and ignored the other buttons.

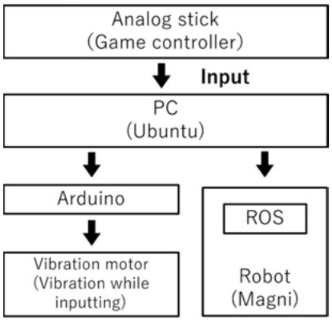

Figure 5: Flow diagram of robot operation.

Figure 6 illustrates the operational environment for moving the robot, featuring the course layout (left) and actual experimental setup (right). The course, designed to be 1000 mm wide, accounts for the increased difficulty posed by the narrower dimensions to inexperienced users. The participants were seated so that they could visually confirm the entire course from start to finish. They faced the direction of the robot's initial straight movement with a course designed for forward motion and a left turn towards the goal.

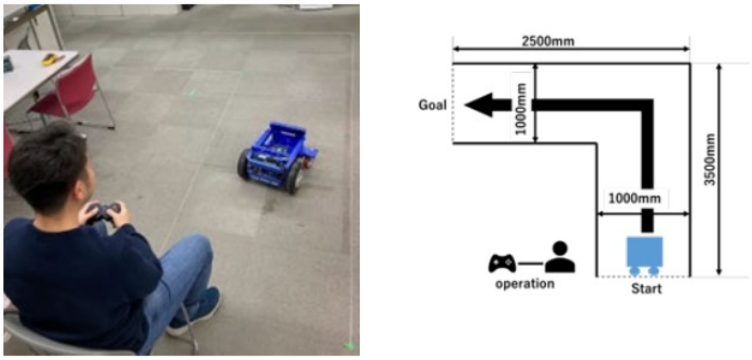

Figure 6: Experimental course (left) and course outline (right).

As shown in Figure 7, to create movements that deviated from intuition, we rotated the input axis of the analog stick by 20 °to the left. This change pushed the stick downward and tilted it slightly to the right. This adjustment, which involved a rotation matrix formula applied to the software, was performed to intentionally alter the stick axis for a specific experimental effect.

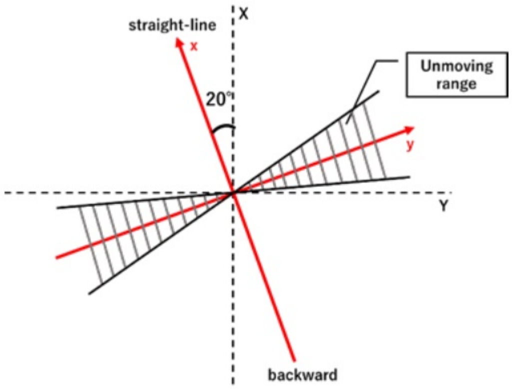

Figure 7: Analog-stick operating range

$$X = x\cos\theta + y\sin\theta \tag{1}$$

$$Y = -x\sin\theta + y\cos\theta \tag{2}$$

The choice of 20° was determined during prototyping. A greater tilt would drastically alter the interface, deviating from our initial design discussions, while a smaller angle would not differ sufficiently from intuitive expectations, considering potential controller inaccuracies. Forward or right turns were enabled by tilting the analog stick along the red line. The inputs in the diagonally marked area were designed to halt the robot and prevent abnormal turning behaviors. This nonmoving zone was defined to avoid unexpected actions during turns, with explanations provided to Group A, which expected movements as anticipated. However, this information was withheld for Group B, which experienced unexpected movements. The vibration feedback was activated for forward and backward inputs along the red axis. Despite issues such as tire orientation and axle alignment causing unintended reactions during straight movements, this has been clarified previously. Both groups were informed that the sensor (vibration) would be activated when the robot moved forward upon direct stick input.

The study participants were 12 individuals in their twenties who were evenly divided into Groups A and B, with each group containing six members. A Likert-scale questionnaire was used to gather the responses. The experimental conditions are listed in Table 2.

Table 2. Questionnaire items.

Questions	Evaluation
Visual	
Q1. How much did your experience and intuition influence your decision-making?	1 : Not affected at all - 7 : Highly affected
Q2. How confident did you feel operating the robot?	1 : Not reliable at all - 7 : Highly reliable
Q3. I can't rely on this interface.	1 : Don't agree at all -7 : I agree very much.
Visual and vibration	
Q4. How much did your experience and intuition influence your decision-making?	1 : Not affected at all - 7 : Highly affected
Q5. How confident did you feel operating the robot?	1 : Not reliable at all - 7 : Highly reliable
Q6. I can't rely on this interface.	1 : Don't agree at all - 7 : I agree very much.
Q7. How much did you rely on your "experience and intuition," " visual," and "vibration," respectively, to perform the operation?	1 : Not affected at all -7 : Highly affected

RESULTS

This section details the experimental data using box plots that visually represent the distribution and central tendencies of the data. In the box plots, the line inside the box denotes the median, providing a measure of the central tendency, and the "×" mark signifies the mean, offering insights into the overall average of the data. Given the limited number of data points, we employed nonparametric tests, specifically the Wilcoxon rank-sum test and the Wilcoxon signed-rank test, to assess significant differences. Figure 8 summarizes the survey that assessed the impact of individuals' experiences and intuition on their decision-making processes. The results indicated a narrower spread for Group A when relying solely on sight versus when combining sight with vibration. However, with p-values surpassing the significance threshold of 0.05, no significant differences were identified. Figure 9 shows the confidence levels during robot operation, indicating a reduced spread for Group B with the use of sight and sensor (vibration) feedback compared with the other conditions. Conversely, Group A exhibited a larger spread under the same conditions. With p-values exceeding 0.05 across all conditions, this suggests the absence of significant differences.

Figure 10 displays the results of interface reliability. For Group B, the data showed a tighter spread when using only sight as opposed to using sight and vibration together, suggesting higher reliability scores with less dependence on the interface in both the mean and median terms. This time, with a p-value less than 0.05, we observed a statistically significant difference.

Finally, Figure 11 analyses the influence of intuition/experience, sight, and vibration on decision-making. When considering intuition and experience, Group A demonstrated narrower variability with higher mean and median values than Group B, indicating greater consistency within the group. In terms of sight, both groups achieved high scores with limited variability, suggesting uniform responses among participants. Regarding vibration, Group B attained higher mean and median scores, reflecting greater receptivity or sensitivity to this form of feedback. Within each

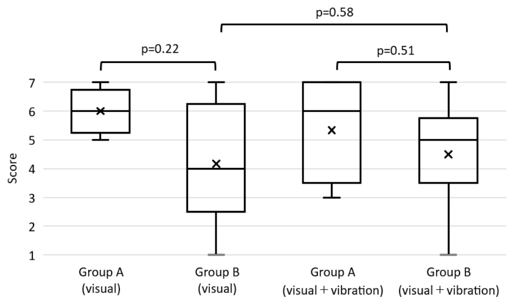

Figure 8: Results of Q1 and Q4 on the degree to which experience and intuition influenced decision-making.

Group A median values ranked from highest to lowest were sight, intuition/experience, and vibration for Group A, and sight, vibration, and intuition/experience for Group B. However, with p-values exceeding 0.05 across all comparisons, this indicates a lack of statistically significant differences.

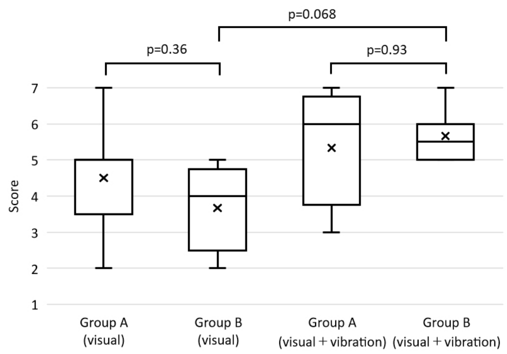

Figure 9: Results of Q2 and Q5 on how confident they felt in operating the robot.

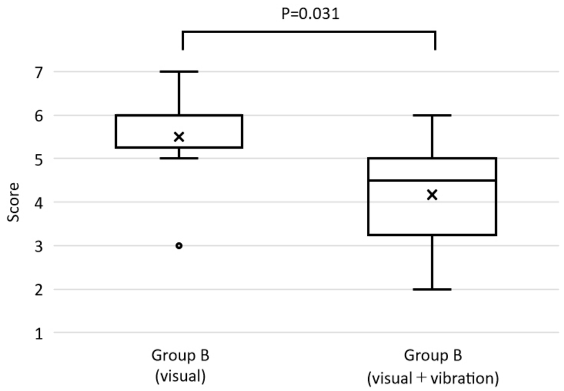

Figure 10: Results for interface reliability for Q3 and Q6.

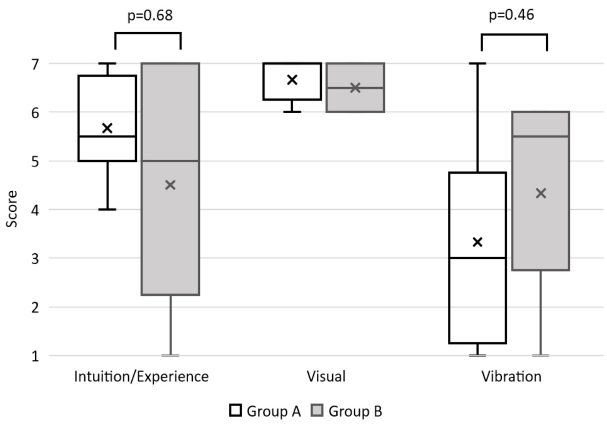

Figure 11: Results for the degree of reliance on "experience and intuition," " visual," and "vibration" in Q7, respectively, to operate.

DISCUSSION

The results, illustrated in Figure 8, evaluate the impact of experience and intuition on decision-making under various conditions. In both the visual and combined visual and vibration (sensor) conditions, Group B displayed lower median and mean values for intuition and experience. This implies that, particularly in critical scenarios, decision-making might deprioritize experience and intuition according to these findings. Despite vibration increasing mean and median values for confidence in robot operation (Figure 9), no significant difference was observed. Thus, it appears that using vibration as an information display method does not directly influence confidence in robot operations. However, the notable increase in the mean and median values for Group B when sensor information was added suggests that supplementing visual observations with sensor data may enhance confidence in robot operation. While comparing solely visual with combined visual and sensor (vibration) inputs in Group B suggests that sensor information might influence confidence, this hypothesis remains unconfirmed. The observed impact can be attributed to various factors, including the frequency of the robot or stick operation. Additionally, we consider that individuals with higher experience levels in robot or stick operations tend to prioritize their experiences and intuition over supplementary information such as vibration. Further research is warranted to explore this aspect thoroughly. As shown in Figure 10, the P-value, related to the analysis of unexpected movement conditions in this experiment, is 0.031, which falls below the established significance threshold of 0.05. This difference was statistically significant. These results suggest that adding vibrational feedback for straight-ahead movements to the interface enhances its reliability in the face of unforeseen robot behaviors. This improvement may stem from feedback that enables operators to verify normal operations through an additional sense beyond visual comprehension. Therefore, displaying information via sensors or other methods on the operational interface can enhance the reliability of operations that lack feedback.

Finally, no significant differences were observed in evaluating reliance on “experience/intuition,” “vision,” and “vibration.” However, the lower mean and median values in unexpected movements suggest a reduced reliance on “intuition” and “experience.” This implies that, in the face of unexpected movements, relying on experience and intuition for assessment becomes challenging, necessitating reliance on additional information. Conversely, the higher mean and median values for vibration in Group B indicate greater openness to incorporating sensor information into decision-making. This likely reflects a willingness among the participants to integrate vibration or sensor information into their decision-making, alongside intuition, experience, and visual information. Visual information was cited in all cases, underscoring that first-hand visual observation is crucial in decision-making when the robot’s behavior is fully observable.

The aforementioned experiments aimed to validate the interface design for enabling humans to respond appropriately during emergencies while using robots. These experiments affirmed that the paramount factor in crafting suitable interfaces for healthcare robots and wearable devices, which is the focus of this special issue, lies in establishing reliability through the integration of the tactile and visual senses.

CONCLUSION

Our study underscores the importance of merging case studies and empirical experiments when designing interfaces to enhance the functionality of service robots, such as caregiving and surgical robots. Traditionally, focusing on case studies or empirical experiments in isolation, combining these yields deeper insights. By focusing on 'mistakes in human assumptions' and behavior in unexpected scenarios, our study shows integrating visual and tactile feedback significantly boosts human-robot interaction reliability. This holistic approach is essential for developing more intuitive and effective robots that minimize human error in critical environments. This study is limited to engineering students with a baseline understanding of robots. Nevertheless, the proposed approach could be applied to a broader user base. For future research, it is crucial to conduct experiments with participants with diverse experience levels, not just those well-versed in robotics or analog stick use. Adopting this inclusive approach will enhance our comprehension of HRI dynamics, ensuring that robot interfaces meet the varied real-world needs of users in both caregiving and surgical contexts.

ACKNOWLEDGMENT

The authors express their gratitude to Yurai Horikoshi and Shintaro Suginoma for their invaluable technical assistance during the experiments.

REFERENCES

A. Anokhin and E. Alontseva, "Implementation of human–machine interface design principles to prevent errors committed by NPP operators," Advances in Intelligent Systems and Computing, pp. 743–753, Aug. 2018.

C.-M. Chang, K. Toda, X. Gui, S. H. Seo, and T. Igarashi, "Can Eyes on a Car Reduce Traffic Accidents?," Sep. 2022.

D. A. Norman, The Psychology of Everyday Things. New York, Ny: Basic Books, 1988.

D. Miller et al., "Behavioral Measurement of Trust in Automation," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 60, no. 1, pp. 1849–1853, Sep. 2016.

E. Awad et al., "The Moral Machine experiment," Nature, vol. 563, no. 7729, pp. 59–64, Oct. 2018.

E. Dasho, L. Kuneshka, and E. Toci, "Information Technology in Health-Care Systems and Primary Health Care," Open Access Macedonian Journal of Medical Sciences, vol. 10, no. E, pp. 1919–1926, Dec. 2022.

F. M. F. Verberne, J. Ham, and C. J. H. Midden, "Trusting a Virtual Driver That Looks, Acts, and Thinks Like You," Human Factors: The Journal of the Human Factors and Ergonomics Society, vol. 57, no. 5, pp. 895–909, Apr. 2015.

Huseyin Avsar, J. E. Fischer, and T. Rodden, "Designing touch screen user interfaces for future flight deck operations," Nottingham ePrints (University of Nottingham), Sep. 2016.

M. H. Lee, D. P. P. Siewiorek, A. Smailagic, A. Bernardino, and S. B. Bermúdez i Badia, "A Human-AI Collaborative Approach for Clinical Decision Making on Rehabilitation Assessment," Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, May 2021.

M. H. MoghimiEsfandabadi, M. H. Djavareshkian, and S. Abedi, "Significance of Aviation Safety: Its Evaluation and Ways to Strengthen Security," in International Journal of Reliability, Risk and Safety: Theory and Application, vol. 6, no. 2, pp. 37–45, 2023.

M. Nakamura, K. Okajima, Y. Matsumoto, T. Tanaka, K. Iijima, and M. Nihei, "Effectiveness of Continuous Grip Strength Measurement Using Social Assistive Robots on Older Adults at Home," J. Robot. Mechatron, vol. 33, no. 4, pp. 719–729, Aug. 2021.

O. Asan, A. E. Bayrak, and A. Choudhury, "Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians," Journal of Medical Internet Research, vol. 22, no. 6, p. e15154, Jun. 2020.

R. Moura, M. Beer, E. Patelli, J. Lewis, and F. Knoll, "Human error analysis: Review of past accidents and implications for improving robustness of system design," in Proceedings of the European Safety and Reliability Conference (ESREL) 2014, Wroclaw, Poland, pp. 37–45, September 2014.

Riku Tsunori and Y. Okada, "Issues of Human Error Management in Introducing New Technology for Bridge Inspection Drones," Advances in intelligent systems and computing, pp. 238–245, Jun. 2019.

S. K. Sahoo and B. B. Choudhury, "Challenges and opportunities for enhanced patient care with mobile robots in healthcare," Journal of Mechatronics and Artificial Intelligence in Engineering, Aug. 2023.

S. Miura et al., "Evaluation of Hand-Eye Coordination Based on Brain Activity," Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 19, no. 1, pp. 143–151, Jan. 2015.

S. W. Jensen, J. L. Salmon, and M. D. Killpack, "Trends in Haptic Communication of Human-Human Dyads: Toward Natural Human-Robot Co-manipulation," Frontiers in Neurorobotics, vol. 15, Feb. 2021.

T. Matsuo, "医事故防止に心理学はどのように献できるか - オミッションエラを防止する -," presented at the 第73回日本心理学会大会, 敬学 214, Aug. 2009, in Japanese.

T. Takano, A. Ikeda, I. Abe, and T. Kikuchi, "Improvement of Haptic Interface for Teleoperation Endoscopic Surgery Simulators Using Magnetorheological Fluid Devices," J. Robot. Mechatron, vol. 34, no. 6, pp. 1268–1276, Dec. 2022.

Teppei Tsujita et al., "Development of Haptic Interface for Neurosurgical Simulators with Micro Scissors Module for Displaying the Cutting Force," J. Robot. Mechatron, vol. 34, no. 6, pp. 1284–1296, Dec. 2022.

W.-F. Hsieh, E. Sato-Shimokawara, and T. Yamaguchi, "Investigation of Robot Expression Style in Human-Robot Interaction," J. Robot. Mechatron, vol. 32, no. 1, pp. 224–235, Feb. 2020.

X. Luo, Q. Liu, and Z. Qiu, "A Correlation Analysis of Construction Site Fall Accidents Based on Text Mining," Frontiers in Built Environment, vol. 7, Jun. 2021.

Y. Yamazaki, M. Ishii, T. Ito, and T. Hashimoto, "Frailty Care Robot for Elderly and its Application for Physical and Psychological Support," Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 25, no. 6, pp. 944–952, Nov. 2021.

Z. Njolomole and R. Kalimuthu, "Artificial Intelligence IoT and their Future on Aviation Safety," Journal of Information Technology & Software Engineering, vol. 11, no. S5, Article no. 002, Oct. 2021.