Investigating the Cognition-Control Pattern of Multi-Worker Human-Robot Collaboration in Construction

Di Liu[1] and Youngjib Ham, Ph.D.[2]

[1]Ph.D. Student, Dept. of Construction Science, Texas A&M Univ., College Station, TX. Email: catsquito@tamu.edu
[2]History Maker Homes Endowed Associate Professor, Dept. of Construction Science, Texas A&M Univ., College Station, TX (corresponding author). Email: yham@tamu.edu

ABSTRACT

When operating a construction robot (e.g., excavator), the operator's unsafe behavior directly affects the safety risk (e.g., underground utility damage occurrence during excavation process). Operators' behavior is greatly influenced by the surrounding environment and further the communication with other coworkers (i.e., spotter), and thus there is a need for studying human factors during work by investigating how the operator-spotter interaction affects the operator when performing the task in a challenging work environment. In this paper, we investigate how the different levels of environmental complexities and operator-spotter communication channels affect operators' performance and the accident occurrence during work. A human-centered experiment is designed and conducted in environmentally realistic scenarios based on immersive virtual reality. The task of operating a virtual excavator as well as the interaction between the operator and a spotter are performed in realistic jobsites in which a series of environmental stimuli are modeled and simulated. Operators' cognitive responses and work performance are assessed by subjective evaluations and instrument-based measurements (i.e., eye-tracking). This study can establish a better understanding about the effectiveness of between-worker communication and worker-to-robot interaction during robot operation as well as the influence of environmental visual and auditory stimuli on the teleoperator.

INTRODUCTION

Excavator, as one of the most popular construction robots, is a major source for construction accidents which accounts for the largest proportion of fatalities on site. It causes great damage to human life and the economy. One of the most fatal excavation accidents was the collision with underground utility lines. In the U.S., the number of such incidents is between 400,000 and 800,000 per year. This causes underground utility damage represents a major drag on national economies, $50-100 billion annually in the U.S. Unfortunately, this type of accident is hard to avoid, especially in challenging urban areas where are full of unpredictable environmental factors. It is well known that a construction jobsite in a crowded urban environment often has a dynamic and unstructured nature. This caused great challenges to construction work performance and safety. Some of these unpredictable challenging factors in an urban jobsite are task-related factors, including inaccurate information of underground utility locations, and uncomfortable workspace for workers. These directly affect performance and safety. Other factors are environmental distractors including surrounding people or activities, visual distractors such as traffic or urban facilities, environment noise.

To prevent underground utility damage, the most common practice is to use 811, "Call-before-Dig", a nationwide phone call or online service to get the location in advance. However, this system is not always sufficient due to unreliable location information caused by inaccurate or lost marks or discrepancy between as-built data in drawings and the actual location, especially in crowded urban areas. Advanced techniques such as GPR or LiDar and other remote sensing and detecting methods improve the location information yet require a well-trained geo-technician. Nevertheless, despite the advantages of existing practices, more than 70% of underground utility strikes happen when utilizing the current damage prevention system, which raises the concern that human-centered practice is overlooked in preventing excavation accidents. Furthermore, in real-life practice, an excavation task often requires multiple construction workers rather than a single operator. This team-based practice can better monitor the entire task environment when the operator concentrated on excavator control. In a team-based excavation practice, a spotter is a human role providing real-time directions to an operator, and an operator is obligated to sense and follow spotter's signals (i.e., verbal commands, body language). Once the excavation starts, using a spotter is a key step to prevent damage. In an online survey conducted among 104 excavator operators who had at least ten hours of experience in operating an excavator for the last six months, over 85% of participants somewhat or strongly agreed that using a spotter is always necessary for excavation. The spotter oversees excavation tasks and detects potential risks. The collaboration between an operator and a spotter is a common practice in a real-world jobsite.

With this regard, when the operator works with a spotter in a complex urban job site, to ensure the excavation safety and efficiency, there is a necessity to assess the operator's cognitive load and the performance on operating the excavator to complete tasks. The objective of this paper is to investigate how the operator-spotter communication formats and challenging environmental factors affect operator's cognitive load and excavator control performance. To achieve this objective, a virtual excavation environment was designed. To assess the human factors of the excavator operator, dependent variables were measured by instrument-based assessment and subjective assessment respectively.

BACKGROUND

Excavation in the real-life job site is a complex interactive process that involves human-machine interaction, and human-environment interaction (Lee et al. 2022; Lee and Ham 2022). This process is affected by various factors associated with human workers, task, machine, and environment.

First, modeling a close-to-real job environment is a key step for virtual operation, and this process can be challenging and impeded by the algorithm development and hardware issues (Prabhakaran et al. 2022). With limited hardware devices, previous studies used the built-in scenarios of the commercialized simulation platform developed by construction vehicle manufactures and took advantage of the high realistic ready-to-use physics engine (Li et al. 2020; Su et al. 2013; Bhalerao et al. 2017). Customized virtual scenarios are also developed to suit specific research goals (Akyeampong et al. 2014; Ding et al. 2022; Hammad et al. 2016; Hong et al. 2020; Vahdatikhaki et al. 2022; Wallmyr et al. 2019). For instance, to compare the obstacle avoidance performance between different visual interfaces, Hong et al. 2020 and Wallmyr et al. 2019 modeled the surrounding environments mimicking the real jobsites including road obstacles, construction accessories, 3D human-character models, and unfavorable weather conditions. Hammad et al. 2016 imported real site data from ArcGIS to Unity3D to simulate the real job location. For the soil excavation task, it is important to simulate real soil physics to achieve the close-to-real experience. Vahdatikhaki et al. 2022 simulated terrain modeling that allowed the virtual soil ground to change dynamically during digging. The real environment is also used for teleoperation, and Mixed Reality (MR) techniques were used to capture the real environmental data. Nevertheless, virtual, or real job environments in these studies are considered as an acceptable yet rather simplified version of real-life scenarios, and these scenarios are insufficient to represent a complicated and dynamic urban job scenario.

Second, regarding human factors, the operator's work performance is the major parameter that reflects how efficiently an operator controls the excavator to complete the task and avoids potential accidents. Task completion time and error rate are the primary measurements (Bhalerao et al. 2017; Feng et al. 2019; Hong et al. 2020; Lorenz et al. 2020; Mavridis et al. 2015; Okishiba et al. 2019) (Ding et al. 2022). For example, Morosi et al. 2019 calculated an overall performance score based on the trajectory of the joints, time, distance, speed, and error. Moreover, operator's cognitive responses directly affect task performance. NASA-Task Load Index (NASA-TLX) is widely used as a standard measurement to subjectively access the mental workload. Instrument-based assessment such as eye tracking techniques have been increasingly implemented in construction hazard detection, visual inspection, and human-machine interaction to understand the attention and mental efforts. Li et al. 2020, 2019 used a mobile eye tracker to assess fatigue in virtual excavator operation. Wallmyr et al. 2019, 2017 used eye tracking to examine the attentiveness on MR display and conventional display in the virtual excavator operation.

Despite the efforts of understanding a single-user (e.g., the operator) excavation in various task scenarios, research gaps were identified. First, there is a dearth of studies to evaluate the operator's task performance and cognitive load associated with the unsafe behaviors stemmed from the challenging environmental factors such as buried utility lines in the task space as well as dynamic environmental distractors. Second, the human factors in the existing studies are examined in a single-user excavation context which is different from team-based real-life practice.

MULTI-WORKER HUMAN-ROBOT COLLABORATION IN CONSTRUCTION

The experiment was conducted in a multi-user hybrid immersive teleoperation system as shown in Figure 1 (Liu et al. 2022; Liu and Ham 2022). A total of 20 subjects were recruited in a random sampling method. A two-by-two factorial human experiment was performed. Two types of communication formats, verbal signal and hand signal, were determined as group factors. All 20 participants were divided into two groups based on the communication formats. Two levels of environments, baseline environment and challenging environment, were determined as condition factors. Specifically, the baseline environment was designed as a roadwork scenario with minimum environment visual and auditory elements. The challenging environment, on the other hand, was designed as a jobsite in a downtown area, including a series of environmental distractors such as crowded urban activities, traffic, and environment noises from people, traffic, and nearby construction works. Prior to the experiment day, participants were required to review an instruction and get familiarized with standard excavator control with joysticks as well as ten signals to direct the movement of major excavator components, including bucket, stick, boom, and cabinet. All signals were represented in verbal and hand formats. To be familiarize with the virtual environment and identity potential discomfort, participants were required to practice the basic excavator operation independently in an immersive virtual environment as well as the operation by following the spotter's signals. The entire practice session lasts 60 minutes. On the experiment day, each participant went through a total of four experimental sessions. In session 1, the participant completed the eye-tracker calibration. In session 2, the participant performed excavation trials with being exposed to two immersive work environments, baseline (BE) vs. challenging (CE), respectively. Each trial lasts 5 minutes approximately. In session 3, excavation trials were repeated in reversed order. A short break was provided between session 2 and 3. In session 4, the participant completed the subjective evaluation.

In each experimental trial, participants were required to excavate full loads of earthwork from the digging area placed in front of the excavator and disposed to the left side. Both excavated and disposal areas were clearly white lined, and the utility were modeled under the ground surface in the excavated area. Participants were required to dig from and dump to the defined task areas as accurately as possible. Moreover, potential collisions with the buried utility lines should be avoided. An avatar that captured the spotter's real-time motions was placed within the operator's field-of-view on the other side of the excavated area. During the entire task trial, participants were required to check and follow the spotter's signals all the time. Experimental data were collected by video recorder, eye-tracker, and a log file attached to the system. Figure 2 shows the structure of task operation and data collection system.

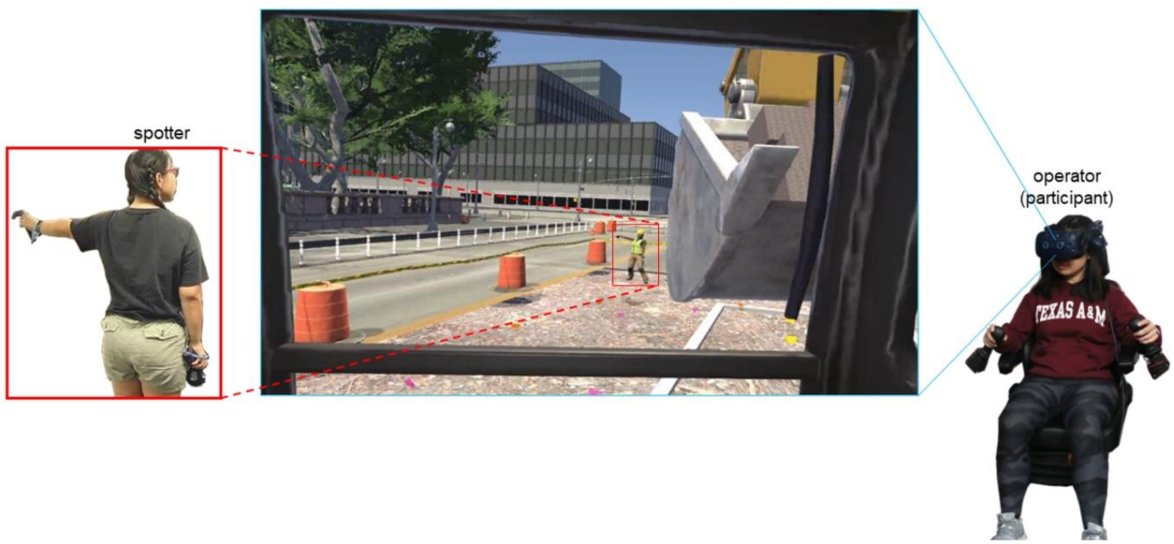

Figure 1. The proposed immersive multi-user communication system

RESULTS AND DISCUSSIONS

NASA-TLX. In this experiment, a modified version of NASA-TLX questionnaire was used. Five dimensions, namely, mental demand, temporal demand, performance, effort, frustration, were included. As shown in Table 1, significant differences are found in all five dimensions in the within-group analysis. This indicates that the type of work environment significantly affects participants' workload assessed by NASA-TLX. Regarding between-group results, the mean value of total score in verbal group (35.8) is higher than in hand group (32.3). According to the results of two-way ANOVA (Table 2), while the environment type has a significant effect ($p < 0.01$, $F = 40.34$), there is no significant effect found regarding the communication formats ($p = 0.43$, $F = 0.63$) or the interaction between two independent variables ($p = 0.41$, $F = 0.71$).

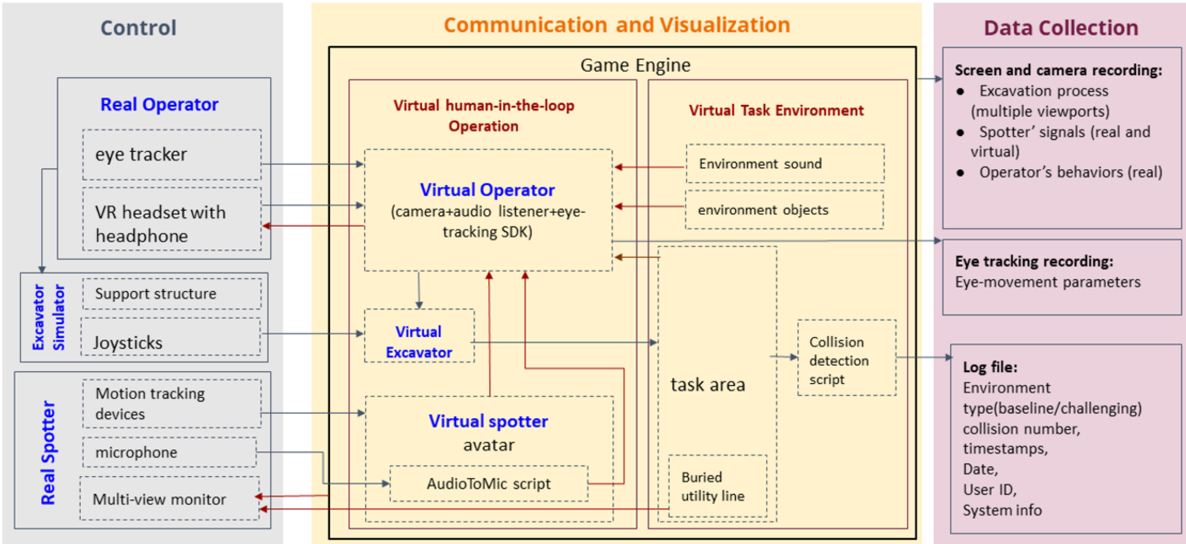

Figure 2. The structure of task operation and data collection system

Table.1. NASA – TLX: two types of environments

	Mental Demand	Temporal Demand	Performance	Effort	Frustration	Total
Statistic p value	29.5 3.73e-06*	97.0 0.0046*	299.0 0.0059*	67.0 0.0003*	21.5 1.07e-06*	29.0 3.88e-06*
*p<0.01						

Table.2. NASA – TLX: Formats x Types

	sum_sq	df	F	PR(>F)
Formats	28.9	1.0	0.630316	4.324e-01
Types	1849.6	1.0	40.340240	2.363e-07*
Formats: Types	32.4	1.0	0.706652	4.061e-01
Residual	1650.6	36.0	-	-
*p<0.01				

Pupillometry. The data of pupil sizes was obtained during all experimental trials. Pupil sizes from both eyes were regressed into mean pupil size. To minimize errors due to excessive sampling, a procedure of smoothing and down sampling was applied. The linear interpolation was applied across blinks. As shown in Figure 3, for the within-group analysis (left), the absolute pupil sizes in the challenging environment are found to be higher than in the baseline environment. For the between-group analysis (right), the absolute pupil size in the hand-signal group and verbal-signal group appears remain similar. In Figure 4, changes of pupil size in within-group remain similar (mean_baseline = 0.08mm, mean_challenging = 0.08mm). For between-group, the mean value of the changes in verbal group is higher (0.11mm) than the mean value in the hand group (0.06mm). Meanwhile, there is no significant difference found in within-group analysis ($p = 0.87$, $F = 0.025$) or between-group analysis ($p = 0.25$, $F = 1.41$).

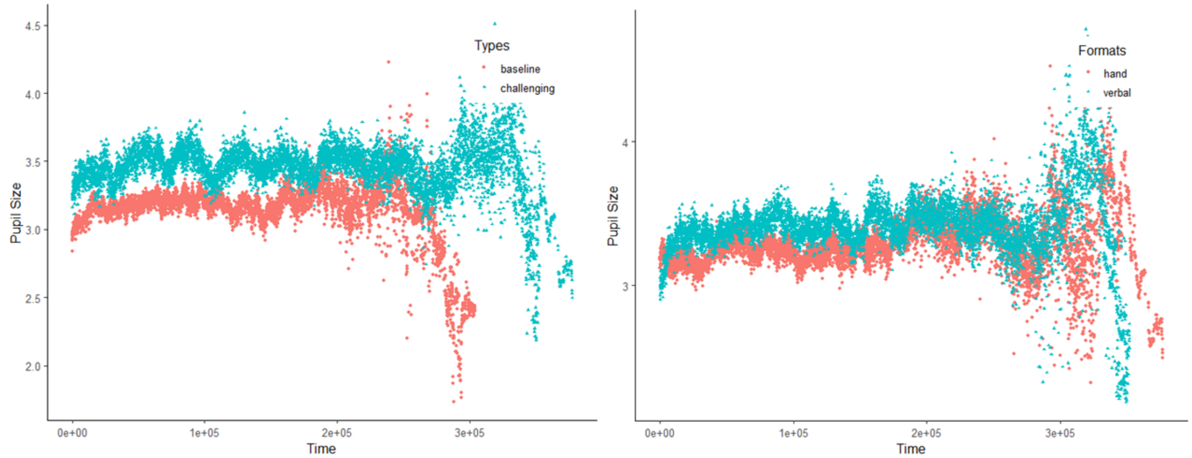

Figure 3. Pupil Size (mm), left – baseline vs challenging (within group), right – hand group vs verbal group (between group)

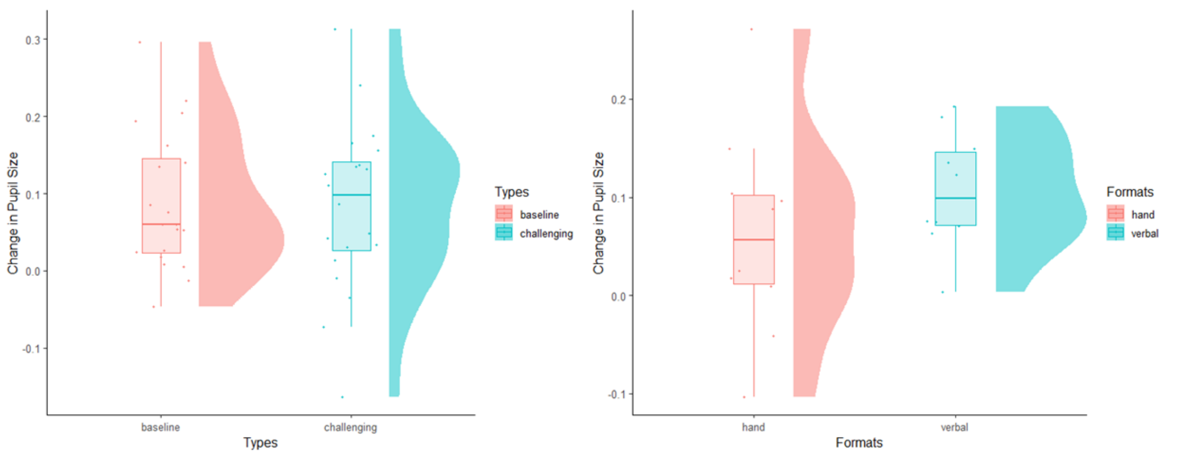

Figure 4. Change in Pupil Size (mm), left – baseline vs challenging (within group), right – hand group vs verbal group (within group)

Performance (number of collisions). The mean value of collisions in a challenging environment (3.4) is higher than the mean value in the baseline environment (0.25). For between group results, the mean value of collision in hand group (1.95) is slightly higher than that of verbal group (1.7). One way ANOVA test shows that the environment types have a significant effect on the number of collisions made ($p < 0.01$, $F = 17.31$). There were no significant effects found in terms of communication formats ($p = 0.74$, $F = 0.10$) or the interaction between two independent variables ($p = 0.33$, $F = 0.98$).

CONCLUSION

This paper explores if the communication formats between the excavator operator and a spotter as well as environmental complexities affect the excavation task performance and cognitive load in a team-based excavation process. A two-by-two factorial experiment with 20 subjects was conducted in a multi-user hybrid-immersive teleoperation system. Subjective assessment of NASA-TLX shows that users' cognitive workload is higher in challenging environments with a significant effect ($p < 0.01$). Pupillometry analysis indicates that cognitive workload is higher in challenging urban jobsites, as well as in the group communicating with verbal signals. The number of collisions also demonstrates that the type of environment significantly affected the safety performance. The contribution of this study is to enable a better understanding on selecting the effective communication strategies based on different environmental complexities in a collaborative excavation work to ensure safety and reduce the mental workload. Future works lies on an in-depth investigation including other assessments such as situational awareness and attention.

ACKNOWLEDGEMENT

This material is in part based upon work supported by the National Science Foundation (NSF) under Award #2026574. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

Akyeampong, J., S. Udoka, G. Caruso, and M. Bordegoni. 2014. "Evaluation of hydraulic excavator Human–Machine Interface concepts using NASA TLX." *Int. J. Ind. Ergon.*, 44 (3): 374–382.

Bhalerao, B. N., P. S. Dunston, and R. W. Proctor. 2017. "Use of PC-based Simulators to Train Basic Control Functions of a Hydraulic Excavator: Audiovisual Instruction Contrasted with Hands-On Exploration." *Int. J. Human–Computer Interact.*, 33 (1): 66–74. Taylor & Francis.

Ding, R., M. Cheng, Z. Han, F. Wang, and B. Xu. 2022. "Human-machine interface for a master-slave hydraulic manipulator with vision enhancement and auditory feedback." *Autom. Constr.*, 136: 104145.

Feng, H., C. Yin, R. Li, W. Ma, H. Yu, D. Cao, and J. Zhou. 2019. "Flexible virtual fixtures for human-excavator cooperative system." *Autom. Constr.*, 106: 102897.

Hammad, A., F. Vahdatikhaki, and K. El-Ammari. 2016. *Conceptual Framework of Training Simulator for Heavy Construction Equipment Integrating Sensory Data, Actual Spatial Model, and Multi-Agent System*. 8.

Hong, Z., Q. Zhang, X. Su, and H. Zhang. 2020. "Effect of virtual annotation on performance of construction equipment teleoperation under adverse visual conditions." *Autom. Constr.*, 118: 103296.

Lee, J. S., and Y. Ham. 2022. *Exploring Human-Machine Interfaces for Teleoperation of Excavator*. 757–765. American Society of Civil Engineers.

Lee, J. S., Y. Ham, H. Park, and J. Kim. 2022. "Challenges, tasks, and opportunities in teleoperation of excavator toward human-in-the-loop construction automation." *Autom. Constr.*, 135: 104119.

Liu, D., and Y. Ham. 2022. "Multi-User Immersive Operation and Communication In Virtual Reality: A Research Case On Robotic Excavator." *CONVR 2022 Proc.* Seoul, Korea.

Liu, D., Y. Ham, J. Kim, and H. Park. 2022. "Towards a Collaborative Future in Construction Robotics: A Human-centered Study in a Multi-user Immersive Operation and Communication System for Excavation." *Proc. 1st Future Constr. Workshop Int. Conf. Robot. Autom. ICRA 2022. International Association for Automation and Robotics in Construction (IAARC)*.

Lorenz, S., J. R. Helmert, R. Anders, C. Wölfel, and J. Krzywinski. 2020. "UUX Evaluation of a Digitally Advanced Human–Machine Interface for Excavators." *Multimodal Technol. Interact.*, 4 (3): 57. Multidisciplinary Digital Publishing Institute.

Mavridis, N., G. Pierris, P. Gallina, Z. Papamitsiou, and U. Saad. 2015. "On the subjective difficulty of Joystick-based robot arm teleoperation with auditory feedback." *2015 IEEE 8th GCC Conf. Exhib.*, 1–6.

Okishiba, S., R. Fukui, M. Takagi, H. Azumi, S. Warisawa, R. Togashi, H. Kitaoka, and T. Ooi. 2019. "Tablet interface for direct vision teleoperation of an excavator for urban construction work." *Autom. Constr.*, 102: 17–26.

Prabhakaran, A., A.-M. Mahamadu, and L. Mahdjoubi. 2022. "Understanding the challenges of immersive technology use in the architecture and construction industry: A systematic review." *Autom. Constr.*, 137: 104228.

Vahdatikhaki, F., A. K. Langroodi, L. olde Scholtenhuis, and A. Dorée. 2022. "Feedback support system for training of excavator operators." *Autom. Constr.*, 136: 104188.

Wallmyr, M., T. A. Sitompul, T. Holstein, and R. Lindell. 2019. "Evaluating Mixed Reality Notifications to Support Excavator Operator Awareness." *Hum.-Comput. Interact. – INTERACT 2019, Lecture Notes in Computer Science*, D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris, eds., 743–762. Cham: Springer International Publishing.