A USDOT University Transportation Center

New York University

Rutgers University

University of Washington

The University of Texas at El Paso

City College of New York

Work Zone Safety III: Calibration of Safety Notifications through Reinforcement Learning and Eye Tracking

May 2022

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2.Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Work Zone Safety III: Calibration of Safety Notifications through Reinforcement Learning and Eye Tracking		5. Report Date May 2022 6. Performing Organization Code
7. Author(s) Semiha Ergan, Kaan Ozbay, Suzana Duran Bernardes, Daniel Lu, Fan Zuo		8. Performing Organization Report No.
9. Performing Organization Name and Address Connected Cities for Smart Mobility towards Accessible and Resilient Transportation Center (C2SMART), 6 Metrotech Center, 4th Floor, NYU Tandon School of Engineering, Brooklyn, NY, 11201, United States		10. Work Unit No. 11. Contract or Grant No. 69A3551747119
12. Sponsoring Agency Name and Address Office of the Assistant Secretary for Research and Technology University Transportation Centers Program U.S. Department of Transportation Washington, DC 20590		13. Type of Report and Period Covered Final Report, 3/1/2021-5/2/2022 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Despite increased regulations, restrictive measures, and devices used for warnings, work zone injuries and fatalities are still observed at highway construction projects with alarms/notifications being ignored. With a vision to reduce the number of injuries and fatalities, Phase 3 of the research team's worker safety project extends the original scope and adds two new main components, including the addition of eye-tracking for identifying worker attention under dangerous situations and a reinforcement learning model used to optimally send alarms to workers to maximize their attentions along with wide deployment and demonstration of the team's previous C2Smart research effort. This phase of the project aims to make sense of the biometric sensor data (i.e., heart rate and pupil movements while workers omit or accept safety notifications) through state-of-the-art reinforcement learning approaches. The outcomes of this research will bring an understanding to the unknowns of worker behaviors on why they decide to ignore/accept notifications for calibration of when and at what frequency to send notifications to workers for a better acceptance rate. Key questions this research answers are, at what conditions workers ignore/response to warnings at work zones? How we can calibrate notification systems for getting responsive actions from workers? What are the modalities, frequencies, and timings of pushing notifications in these calibrated systems? Through wearable sensors, hardware integrated realistic representations of work zones in virtual reality and eye tracking, in this phase of the project the team will widely have pilot demonstrations of the integrated platform to collect worker behavioral and biometric (heart rate, eye-tracking) responses to alarms/warnings/notifications issued under realistic scenarios and modalities of warning mechanisms (e.g., sensory, visual, audial) that were developed in earlier phases of this project, and mine these captured data towards understanding human behaviors in response to modalities of notifications.

17. Key Words	18. Distribution Statement Public Access		
19. Security Classif (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No of Pages 55	22. Price

Form DOT F 1700.7 (8-69)

Work Zone Safety III: Calibration of Safety Notifications through Reinforcement Learning and Eye Tracking

Dr. Semiha Ergan
New York University
ORC-ID 0000-0003-0496-7019

Dr. Kaan Ozbay
New York University
0000-0001-7909-6532

Suzana Duran Bernardes
New York University
0000-0002-3012-0631

Daniel Lu
New York University
0000-0001-5198-7878

Fan Zuo
New York University

C2SMART Center is a USDOT Tier 1 University Transportation Center taking on some of today's most pressing urban mobility challenges. Some of the areas C2SMART focuses on include:

Urban Mobility and Connected Citizens

Urban Analytics for Smart Cities

Resilient, Smart, & Secure Infrastructure

Disruptive Technologies and their impacts on transportation systems. Our aim is to develop innovative solutions to accelerate technology transfer from the research phase to the real world.

Unconventional Big Data Applications from field tests and non-traditional sensing technologies for decision-makers to address a wide range of urban mobility problems with the best information available.

Impactful Engagement overcoming institutional barriers to innovation to hear and meet the needs of city and state stakeholders, including government agencies, policy makers, the private sector, non-profit organizations, and entrepreneurs.

Forward-thinking Training and Development dedicated to training the workforce of tomorrow to deal with new mobility problems in ways that are not covered in existing transportation curricula.

Led by New York University's Tandon School of Engineering, **C2SMART** is a consortium of leading research universities, including Rutgers

Disclaimer

Acknowledgements

The authors would like to acknowledge the funding agencies for this project: C2SMART funding under the grant number (69A3551747124) and a 50% cost-share by New York University.

Executive Summary

With the goal of reducing the number of injuries and fatalities, this project aims to understand the key parameters (e.g., traffic loads, work zone location characteristics, personal vigilance levels, duration of construction work, safety messages exchanged between workers and installed sensors/vehicle) that play roles in achieving responsive behaviors in workers. Key questions this research will address include (a) what is the level of attention of workers with regard to the current work zone condition (e.g., vehicle head direction/speed toward work zone, worker vicinity to work zone perimeters) when they react to/ignore alarms, (b) what modalities, frequencies, and timings of pushing alarms are most effective, and (c) how notification systems can be calibrated for getting responsive actions from workers.

The first component of this project focused on monitoring construction workers' attention. To that end, we added a new functionality to the current VR platform to track the subjects' attention through his/her head-movement and eye-movement to infer his/her gaze pattern. With the introduction of this method to measure subject's attention, we captured additional critical information about the decision a worker makes (i.e., understanding if the subject was unaware of the dangerous situation or if the subject decided to ignore the alarm after an assessment of the danger level by analyzing their gaze patterns and body movements around the time notifications are sent). The second component focused on creating a calibrated notification delivery mechanism that is optimized to send alarms to grab the maximum attention of workers using reinforcement learning (RL). This RL model relies on the data captured in pilot demonstrations in partnering agencies, with students, staff, and faculty who participated in the IRB approved user studies (IRB FY2020-3946).

The research outcomes showed that on the average it takes 2.5 seconds for a worker to react to an alarm. Data also showed that participants were more responsive to acknowledging alarms that were received with multiple stimuli (vibration + sound vs. only vibration). For additional awareness metrics calculated, workers were quick (between 1.5-3 seconds) to display responses (e.g., head turn, seeing a dangerous car, changing position) in the work zone. Regarding other characteristics of alarms analyzed, it was observed that regardless of duration and repetitions, it was the alarm modality people pay more attention to with vibration and sound combination people being more responsive to for turning towards the direction of traffic. Longer durations (=3.8 seconds as compared to 1.8 seconds) with longer pause periods (i.e., 2 seconds as compared to none or 1 second periods) in alarms receive faster response from participants, and shorter durations with short/no pause periods with less repetitions received faster response for head turns towards the direction of the traffic.

Results also indicated that the RL agent outperforms random alarm actions after a small number of training episodes. The optimization approaches within the field of RL research (PPO, actor-critic, Deep Q-learning) will need to be explored further to see if the RL agent can at least choose alarm actions as expected by Monte Carlo simulations. Improvements to the attention monitoring system, such as distinguishing between when workers see vehicles in their central field of view or just their periphery, can also feed into ways the reward function can vary the final reward value for a better gradient. These improvements to the RL agent reward function, optimization approach, and training configurations will help future work on using RL to optimize alarm characteristics during live VR user studies.

Table of Contents

Work Zone Safety III: Calibration of Safety Notifications through Reinforcement Learning and Eye Tracking i

- Executive Summary v
- Table of Contents vi
- List of Figures viii
- List of Tables ix
- Introduction 1
- Motivation 2
- Background 3
  - Current and Emerging Sensing Technologies Adopted at Traffic Work Zones 3
  - Applications of Virtual Reality in Relation to Worker Safety in the AEC 4
  - Applications of Reinforcement Learning and Optimal Control in Civil Eng. Applications 6
- Research Methodology 7
  - An Overview of the VR-HIL Traffic Simulation Platform 8
  - An Overview of the IRB-approved User Study and Scenarios Implemented 10
  - An overview of the data collected during the user studies 15
  - Types of Alarms Generated via the Smartwatch 20
  - Initial Statistical Analyses on Alarm Characteristics 24
  - Boundary Set for the Reinforcement Learning Model 24
- Results 35
  - Worker Reaction Benchmark Results 35
  - Monte Carlo Reward Simulation 45
  - Reinforcement Learning Results 46
  - Discussion of Findings 49
- Outputs, Outcomes, Impacts 50
  - Outputs 50
  - Outcomes 51
  - Impacts 51
- Conclusion 52
- References 52

List of Figures

Figure 1: Hardware in the loop integration of VR, traffic simulation, worker alarm system, and work zone monitoring system 9
Figure 2: User study setup utilizing the integrated platform 11
Figure 3: A closeup view in virtual world for Scenario 1 12
Figure 4: A closer view from the virtual world showing speeding cars and their detection zones 13
Figure 5: A closer view from the virtual world showing an example collision vehicle 13
Figure 6: A photo from the real experiment boundary tracked with the ultrasonic sensors 14
Figure 7: Illustrations of features that are included in worker attention towards safety alarms ..16
Figure 8: Gaze direction vector of the worker as tracked by the VR headset 17
Figure 9: A sample data from a participant showing the calculated gaze direction angles. Note that the green dots on the plot are indicating head turn speed events, vertical orange lines are indicating triggered events. 18
Figure 10: A sample data showing the calculated number of cars in field of view of a worker. Note that the vertical orange lines are indicating triggered events 18
Figure 11: Screenshot of Unity3D worker's VR perspective (left) and bird's eye view (right). Note the red and blue arrows on the right top corner, which indicate the X-axis and Z-axis location of the worker in the VR environment 19
Figure 12: A sample data from a participant showing the calculated road positions. Note that the green dots on the plot are indicating movement events, vertical orange lines are indicating triggered events 20
Figure 13: A sample data from a participant showing the calculated IBI. Note that the vertical orange lines are indicating triggered events 20
Figure 14: Alarm duration defined relative to the built-in alarms' duration 21
Figure 15: Example showing alarm durations, pause periods with respect to total notification period (with the number of repetitions as 2) 22
Figure 16: Illustrative showing the features of attention that RL agent considers 25
Figure 17: RL parameters defined: actions, observation states, and reward function elements ..26
Figure 18: Illustration of how watch response/acknowledgement time is calculated 28
Figure 19: Illustration of how Vehicle Detection Reaction Time, $\boldsymbol{tveh}$, and Head Turn Reaction Time, $\boldsymbol{thead}$, are calculated 29
Figure 20: Illustration of how Body Movement Reaction Time, $\boldsymbol{tbody}$, is calculated 30
Figure 21: Reward function hierarchy 31

C2SMART
CONNECTED CITIES WITH SMART TRANSPORTATION

Figure 22: Prediction of worker behavior for generation of synthetic data for observation states 34
Figure 23: RL alarm agent training loop....................35
Figure 24: Distribution of reaction times across four response types for all alarm types 37
Figure 25: Histogram of rewards predicted by Monte Carlo Simulation of 1000 random alarm combinations 46
Figure 26: Actions performed by the RL agent over all training episodes. RL agent converged on sound + vibration alarm of mode 6 (duration of 3.8 seconds duration, repeated every second, repeated two times). Red dot: vibration only; Blue dot: Sound + vibration based alarms; Grey dot: RL agent skipped an alarm on a triggered event 47
Figure 27: Moving average rewards over every 100 alarm actions by the RL agent. Blue line: RL actions; Grey line: Random actions....................47

List of Tables

Table 1: Notifications (modes) implemented in user studies 22
Table 2: Preliminary reward and cost values in RL model training 32
Table 3: Statistics for worker reaction times for each response type across all alarms....................37
Table 4: Statistics for watch acknowledgement times in response to alarm configurations (modality, duration, frequency, repetitions) 38
Table 5: Statistics for head turn reaction times in response to each alarm configuration....................40
Table 6: Statistics for vehicle detection times in response to each alarm configuration 41
Table 7: Statistics for body reaction times in response to alarm characteristics....................42
Table 7(continued): Statistics for body reaction times in response to alarm characteristics....................44
Table 8: Statistics for watch acknowledgement times in response to alarm modality 44
Table 9: Monte Carlo simulation based predicted-rewards for 1000 random alarm combinations 45
Table 10: Average reward observed over the first 350 episodes where random alarm characteristics were raised by the RL agent 48

Introduction

In 2019, the Federal Highway Administration reported 762 crashes at roadway work zones, resulting in 842 fatalities, 135 of which were construction workers (BLS, 2019). Nearly a third of these fatalities involved speeding vehicles (FHWA, 2022). A recent survey conducted by the Associated General Contractors (AGC) of America found that sixty percent of highway contractors reported motor vehicle accidents in their construction work zones during 2020 (AGC, 2021).

Current safety measures for workers mainly involve safety training, which generally guides workers to be situationally aware of traffic conditions and alert other workers if they notice any safety hazards. Safety training and traffic construction guidelines offer recommendations for how workers should set up traffic control devices (e.g., cones, barrels) and signage to provide vehicle drivers advanced notice of work zone boundaries, especially for long-term work (i.e., a work with a duration > 3days) and intermediate work (i.e., a work with a duration <= 3 days). Short-term (i.e., a work with a duration <1 hour during daytime), and mobile work zones (i.e., work moves continuously) rely on smaller layouts of traffic control, leaving workers to rely on their own situational awareness and experience to maintain their safety. More novel approaches involve the use of work zone intrusion alarms, which generally use a stationary source of sound and light to warn workers of hazardous vehicles that approach or drive into the construction work zone boundaries (Mishra et al. 2021).

Challenges with developing worker safety alarm systems are two-fold. The first challenge is designing an alarm system with features well suited to ambient work zone conditions, whether it's how loud or bright the alarm sounds / lights are or how it is physically set up on sites. Workers may not hear or see the alarms amidst noisy or dark ambient working conditions. The second challenge is calibrating safety notification characteristics to be most effective in keeping workers alert of safety hazards. This is because workers frequently ignore alarms. This means that any alarm system design needs to emit safety notifications in a way that workers can both perceive and effectively maintain their situational awareness during their work day. Past research involved studies in real-world construction sites, where vehicular accidents are infrequent, or in large scale real-world test beds, where limited accident scenarios are evaluated due to participants' safety. As a more feasible and effective alternative, the research team has developed an integrated platform of virtual reality (VR), micro traffic simulation, and wearable sensors. This platform has been used to evaluate how workers respond to different scenarios of hazardous traffic flows around work zones and identify what alarm characteristics can more effectively evoke worker response and keep their attention to nearby traffic. The platform enabled the research team to collect data on how workers respond to alarms configured with different modalities (e.g., sound and vibration), frequencies, and durations triggered by different intrusion and accident scenarios. The research team has then used this collected data to benchmark worker behavior and physiological responses and train a reinforcement learning (RL) model to optimize these characteristics for increased situational awareness at work zones.

Motivation

The RL technology has found success in addressing various challenges in different domains such as games, flight control systems, and webpage advertising. Although challenges are domain specific, all applications share a need for an intelligent system to explore an unknown solution space and rapidly change its behavior soon after learning more about that solution space. This abstract problem is probably best understood in any complex game, like chess or StarCraft. The best move or action at any point in the game is dependent on many dynamically changing set of factors. The impact of these actions on the outcome of a game may not be understood until after analyzing all the series of events that took place in many game trials and correlating those events to those games' final score. Moreover, the best game player can identify which factors are useful to predict how the game will turn out after making a specific move.

The challenges of worker safety alarm systems particularly mirror the challenges in games and other domains. A construction site, while reasonably controlled with trained workers and site managers, is still a very dynamic environment, with constantly changing conditions such as improvised activities, weather, equipment-related noise, and adjacent traffic flow. Even if an intelligent safety alarm system has enough sensors to be aware of the changing work zone context, the real challenge lies in optimizing the alarm characteristics because 1) no data exists on the optimal characteristics corresponding to a high-dimensional data representation of a worker's behavior and work zone, and 2) the alarm characteristics need to be optimized over a long period of time (e.g., at least multiple work days). This latter challenge has been observed for alarms in general where people experience fatigue after experiencing a number of alarms and frequently ignore them. It is possible that an intense alarm (i.e., high sound volume, high amplitude vibrations) that seems to trigger worker responses in the beginning of a work day may be less effective by the end of the day and will not help contribute to the long-term life safety of workers. With the RL approach, an alarm system can learn which alarm characteristics workers respond to in the long run, while also remain flexible to dynamically changing conditions in real-world work zones. With the research team's previously developed bi-directional integrated VR and micro traffic simulation based platform, we will evaluate how well an RL model's agent maintain worker reactions to its chosen alarm characteristics.

Background

This section provides a synthesis of research work done in relation to (a) current and emerging sensing technologies adopted at traffic work zones, (b) applications of VR in monitoring traffic worker attention, and (c) applications of RL and optimal control in civil applications.

Current and Emerging Sensing Technologies Adopted at Traffic Work Zones

Within the context of this report work zones take place on road/highway construction/ reconstruction projects and share the space with vehicle traffic most of the time. The layout of work zones can vary with the work to be performed (e.g., long term vs. short term), the location of them (i.e., urban vs. rural), and the type of roads (e.g., local roads vs. interstate highways). According to the Manual on Uniform Traffic Control Devices (MUTCD)(USDOT, 2009), work zones can be categorized as mobile, short duration, short-term stationary, intermediate-term stationary, and long-term stationary, depending on the complexity of the work. Mobile work zones are for quick roadwork that takes up to an hour and needs to move intermittently (e.g., pothole filling, surveying, and tree trimming). Short-term work zones refer to the roadwork that happens during daytime for more than one hour but within a single day (e.g., traffic barrier repair, placement of overhead structures, and traffic hardware maintenance/installation) (Tapan et al. 2016). Intermediate-term work zones are for roadwork that happens during daylight for more than a day but up to 3 days or last for more than an hour during nighttime (e.g., pavement markings, barriers, and temporary roadways) (Lewis, 1989). Finally, the long-term work zones are for roadwork that takes more than 3 days (e.g., installation of permanent barriers, marking in long segments of the road). The long-term and intermediate work zones usually have detailed safety guidelines and are thoroughly planned, whereas short-term and mobile work zones have fewer specific safety guidelines and separation from traffic (Wong et al. 2011) Thus, workers in short-term and mobile work zones are more likely to be exposed to the risk of being struck by upcoming traffic (Wong et al. 2011).

The most common safety measures adopted in previous years to improve work zone safety have targeted improving drivers' behaviors. Some of these safety measures aim to control the traffic speed by using fixed or variable message signs, speed display trailers, flagging and lane width reduction (TRB, 2005), and designing work zone layouts based on the Federal Highway Administration (FHWA) regulations (FHWA, 2020). On the other hand, safety measures targeting construction workers focus on the use of intrusion safety alarm systems such as SonoBlaster (i.e., impact-activated intrusion safety alarm system) (Transpo, 2020), Intellicones (i.e., modular radio-based intrusion safety alarm system) (Intellicone, 2020), and Advanced Warning and Risk Evasion Systems (AWARE) (i.e., radar-based intrusion safety alarm system) (Aware, 2020). These intrusion safety alarm systems, along with worker safety training and the use of high-visibility Personal Protective Equipment (PPE) are a few of the measures used to increase work zone safety from the perspective of workers.

Despite current safety measures, workers still find themselves exposed to traffic and at risk of being struck by a vehicle; especially, the workers of mobile and short-term work zones (i.e., work lasts less than one day). Workers like flaggers and surveyors usually have no barriers between them and the traffic, which is one of the reasons why they are at the highest fatality risk (McAvoy et al. 2007; Domenichini, 2017). In addition to having few to no barriers, these road users need to direct their attention to the job they perform, which can limit their visual and auditory response to the events happening in the background. One way of capturing the attention of roadside workers is intrusion alarm systems. Intrusion alarm systems make use of audio and lighting features to safely alarm workers of an intruding vehicle. The primary challenge of these systems is that the auditory and visual cues might get mixed with those present in the work zone. Alternative to existing visual and auditory warning systems, tactile sensory warning systems are another common attention enhancement tool used in work zones, which have been shown to transmit accurate information regarding an intruding vehicle. However, these solutions still need to be further investigated to validate their use and assess their performance in work zones (Sakhakarmi and Park, 2019; Park and Sakhakarmi, 2019), to understand how workers respond to different modalities of safety alarm systems at work zones, without exposing them to real danger- hence the scope of the platform developed in this research work.

Applications of Virtual Reality in Relation to Worker Safety in the AEC

VR has been widely used in construction worker safety studies due to its ability of replicating realistic work zone environments. Furthermore, VR allows for an immersive experience of hazardous situations without actually putting workers in physical harm. Studies using VR in the domain of worker safety concentrates on three main aspects, as safety training and education, safety planning, and safety inspection (Li et al. 2018). Despite the wide adoption of VR in construction worker safety applications, challenges remain for horizontal work zones, as the traffic pattern represented in these VR environments are often static and follow pre-determined trajectories. This prevents workers from experiencing a realistic work zone environment, and more importantly, limits the possibility of studying the impact of worker behavior on traffic patterns. There exist research efforts in integrating traffic simulations with VR. However, previous studies implementing VR and traffic simulation in the context of safety focus on the perspective of the driver. For example, Bella (2005) performed calibration and validation of a driving simulator to study the effects of temporary traffic signals on the traffic speeds in different areas of a work zone. The variation in drivers' speeding behavior in the vicinity of work zones under different scenarios continued to be a subject of interest in recent studies (McAvoy et al. 2007; Domenichini, 2017). Another use of VR for work zone safety is for the analysis of key factors contributing to work zone crashes from the perspective of drivers (MacAvoy et al. 2007). Even though traffic simulation has been integrated with VR in work zone safety studies to simulate traffic patterns in VR environments, there is a gap in the literature regarding its use to understand the behavior in work zones of construction and road workers, as well as dynamic representation of traffic patterns under the influence of construction worker activities.

For safety training and education, traditional means (e.g., lectures, videos, and demonstrations) suffer from low engagement and learning from trainees. VR-based safety training aims to simulate realistic work zones where trainees can rehearse tasks safely while identifying potential risk factors for an operation (Lukas et al. 2008). The realistic and interactive VR environments increase workers' attention and enthusiasm to learn and to improve safety knowledge. Detailed information of the construction project, such as site layout, egress access, and material locations can be replicated in VR, which allows trainees to better understand the construction operations planned onsite. Studies comparing VR-based safety training with traditional methods concluded that VR is a more efficient and engaging platform for trainees to learn safety related knowledge (Burke et al. 2011). On the other hand, VR-based training is found to be superior to training in the physical work zone environment due to its ability to simulate unsafe scenarios without putting trainees in danger. Furthermore, studies comparing VR-based training and training in physical settings concluded that VR requires fewer mental efforts from the trainees since demonstrations onsite can be overwhelming for trainees (Lin et al. 2011).

Safety planning refers to the identification of potentially unsafe scenarios and practices in a construction project prior to the actual construction. Traditional safety planning relies on 2D drawings, accident reports, and computer-aided design models (CAD files). This prevents construction crew, site safety managers, and owners from intuitively understanding the site layout, design requirements and previous incidents' circumstances (Pearlman, 2014). In this regard, VR-based safety planning offers superior level of immersion as compared to traditional mediums of planning documents (Du et al. 2018). As a result, construction crews can achieve higher ability of risk assessment and level of situation awareness, which are critical to site safety (Zuluaga, 2016). Evidence showed that in VR environments, construction workers are more likely to identify unsafe practices and scenarios as compared to photos and videos (Zuluaga, 2016).

Safety inspections is the critical examination of the construction sites in terms of site safety and worker behavior (Gheisari and Esmaeili, 2019). To effectively monitor worker behavior and site environment, site safety managers do visual inspections during walk-arounds or through the help of cameras installed at various locations onsite. However, these current practices have drawbacks, as physical walk- arounds take time and prevent the site safety manager to holistically understand the overall safety condition of the job site. On the other hand, video streams from static camera locations provide the flexibility of monitoring multiple locations simultaneously but limit the possibility of freely examine the site environments from different angles (Tuttas et al. 2017). VR-based safety inspection reconstructs the job site in a realistic 3D virtual environment, while locations of worker, material, and equipment can be tracked and represented in VR in real-time, which provide insights of travel speeds, paths, and locations of workers and equipment. Compared to traditional inspection methods such as camera recordings, the VR-based safety inspection enhances the ability of safety managers to identify job site hazards more promptly by enabling more flexible viewing angles and better context-awareness.

This study builds on previous applications of VR in the worker safety domain and adds a traffic simulation component to generate realistic traffic patterns in VR while enables the possibility of studying worker behavior's impact on traffic simulation. This bi-directional communication made possible between VR platforms and simulation modeling tools, as well as the assessment of alternative ways to select among VR and simulation options are the main differences between earlier studies utilizing VR in worker safety studies.

Applications of Reinforcement Learning and Optimal Control in Civil Eng. Applications

Reinforcement learning has found previous success in video games, flight controls, webpage advertising, and other domains where the best action to take within a dynamic environment is not readily available due to a lack of available data and near infinite possible combinations of environment features (Vinyals et al. 2019; Ng et al. 2006; Tang et al. 2013; Yang and Lu, 2016). RL algorithms therefore attempt to learn through trial and error, recording the actions taken, the outcomes of those actions, and the state of the environment features it observes over time. In effect, these algorithms can improve their performance as they collect more data in real-time, rather than wait for a large dataset to be populated beforehand. A particular early example of RL, multi-arm contextual bandits, found particular success in identifying, among a large set of online ad images, the one ad that most likely prompts users to click on them in response (Tang et al. 2013; Yang and Lu, 2016). As their name suggests, multi-arm bandits modify a set of variables and optimize them in response to an immediately observed payoff or reward, like an octopus playing multiple slot machines. Contextual bandits also factor in information and constraints regarding the surrounding environment (e.g., maximum ad image size allowed, surrounding webpage content). In our study, the objective of identifying the alarm characteristics based on worker reactions is quite similar to the multi-arm contextual bandit problem and motivates us to investigate RL approaches for calibrating alarm characteristics. More notable and sophisticated RL approaches attempt to optimize a series of actions after observing how those actions affect a final result. These approaches have demonstrated significant success in games (Schrittwiser et al. 2020). Policy Proximation Optimization (PPO) has shown promise within the category of policy gradient algorithms, which attempt to find the best actions based on predictions of how rewards may change given a slight change in policy (i.e., stochastic gradient ascent) (Schulman et al. 2017). Our initial RL model utilizes PPO to select the best alarm characteristics over the course of a single VR user trial.

Based on investigations in control systems research, RL has also been explored in the transportation systems, construction engineering, and facilities management domains for optimal control of HVAC systems, construction robotics, and traffic controls (Zou et al. 2020; Lee and Kim, 2021; Santos et al. 2013; Wu et al. 2022). Given the potential safety risks of deploying RL-based control systems to learn from real-world environments, RL research in AEC/FM relies on specialized simulation environments for training and evaluating the RL models (e.g., building energy simulation, construction site, roadway networks). These simulations must replicate the physics of the real-world environment while being able to generate a variety of possible scenarios the control systems could perform under. In effect, these simulation environments have a means of generating synthetic training data for the RL models to learn from. Our initial RL model utilizes a simulation environment consisting of SUMO for simulating the same traffic conditions faced by users in VR studies, as well as a separate model for replicating how those users physically behaved (e.g., head turns, body movement, tapping the smartwatch). Both components of the simulation environment are necessary for the RL model to freely select different alarm characteristics, observe different possible outcomes, and optimize those characteristics accordingly.

Novel applications for RL in worker safety have been somewhat limited, in part because of a limited availability of data on construction workers and means for simulating how workers behave in construction sites. Recently developed cost effective VR systems enable researchers to observe how workers could behave in a variety of construction scenarios while easily recording data on their instantaneous movements and physiology. One recent study has started to investigate using VR and RL to observe when workers are in an unsafe body posture and how interactive displays can alert workers from moving into those poses (Akanmu et al. 2020). The proposed VR system has yet to yield results on the performance of RL based approaches for posture alerts. Overall, given the prior success of RL applied in both AEC and outside domains, there is a need to evaluate how such optimization approaches can perform to improve alarms for worker safety in uncertain and dynamic roadway construction sites.

Research Methodology

The research methodology included developing a bidirectional VR and micro traffic simulation based platform, using this platform for capturing human behavioral data in dangerous traffic scenarios through biometric sensors and smart watch applications, and using these data to statistically analyze user behaviors as well as building RL agents to eliminate unnecessary alarms and calibrate alarms based on alarm characteristics (modality, frequency, duration of alarms) that people are more responsive to.

Hence, in this section we provide an overview of the (a) previously developed VR and micro traffic simulation integrated platform, (b) IRB approved user study to collect human behavior data in dangerous traffic scenarios, and scenarios implemented in VR, (c) data collected during the user studies, (d) type of alarms generated via smart watches, and (e) initial statistical analyses on alarm characteristics, and (f) boundaries of the RL model. Findings of the research has been presented in the results section.

An Overview of the VR-HIL Traffic Simulation Platform

Previous year's work by the research team resulted in the development of an experiment platform integrating virtual reality (VR) and hardware-in-the-loop (HIL) micro-traffic simulations for testing and evaluating characteristics of alarms that alert construction workers of safety hazards. This integrated platform enables hardware-in-the-loop for synchronous VR, traffic simulation and sensor interactions. It allows a two-way information flow between the virtual work zone and the worker safety system, with the help of an application server to relay the information. As seen in Figure 1, there are two components of the platform, as (1) the virtual work zone, which includes a traffic simulation tool, a proxy server, and a VR environment, where traffic patterns are realistically simulated in VR and worker behaviors in VR circle back to the simulation; and (2) the worker safety system, which includes monitoring hardware and alarming hardware, where the worker location with regard to the work zone perimeter was monitored by sensors and the worker receives alarms whenever a dangerous situation was detected in the VR environment. The connection between sensing (i.e., ultrasonic sensor) and VR only initiates an alarm to be sent to the smartwatch worn by the user, rather than dictating the traffic movements in the VR environment directly. These two parts are connected through application servers (i.e., simulation application server, and VR application server) to relay information. The combination of these three components generates realistic traffic flow structured upon reliable and complex traffic models to be presented in the VR environments and allows for a platform that can be used to test worker safety hardware by conducting user studies in VR.

Essentially, the platform can be simplified as an implementation of the information flow pathways depicted in Figure 1, including a feedforward path, a feedback path and a HIL path. The feedforward path transmits the vehicular and traffic control information simulated in a traffic simulation tool to the VR environments. Reversely, the feedback path brings the traffic control changes that occurred in the VR environments back to the traffic simulation tool. The HIL path bridges sensor hardware deployed in the user studies to monitor worker location (i.e., ultrasonic sensors) and alarm workers of potentially dangerous situations (i.e., a smartwatch) and the virtual work zone, where simulated traffic pattern can cause hazardous situations, and traffic simulation can be impacted by worker behaviors (e.g., step out of work zone boundaries). In the following subsections, each component of the platform along with possible alternatives will be introduced in detail, each alternative will be evaluated based on the challenges identified, and implementation details using the selected alternatives for each component will be provided.

This platform utilized application programming interfaces (APIs) that allowed micro traffic simulation software (e.g., Simulation of Urban Mobility -SUMO) to control vehicle trajectories in VR environments (e.g., Unity3D). Workers in the VR environment could perform virtual work tasks that could feedback to traffic simulation vehicle trajectories in real-time. For instance, if a worker placed cones in the path of traffic in Unity VR, those actions would be tracked back to the SUMO, which would simulate vehicles moving to adjacent lanes to avoid the cones. The feedback loop between traffic simulation and VR environments models a virtual work zone for testing worker safety hazard scenarios.

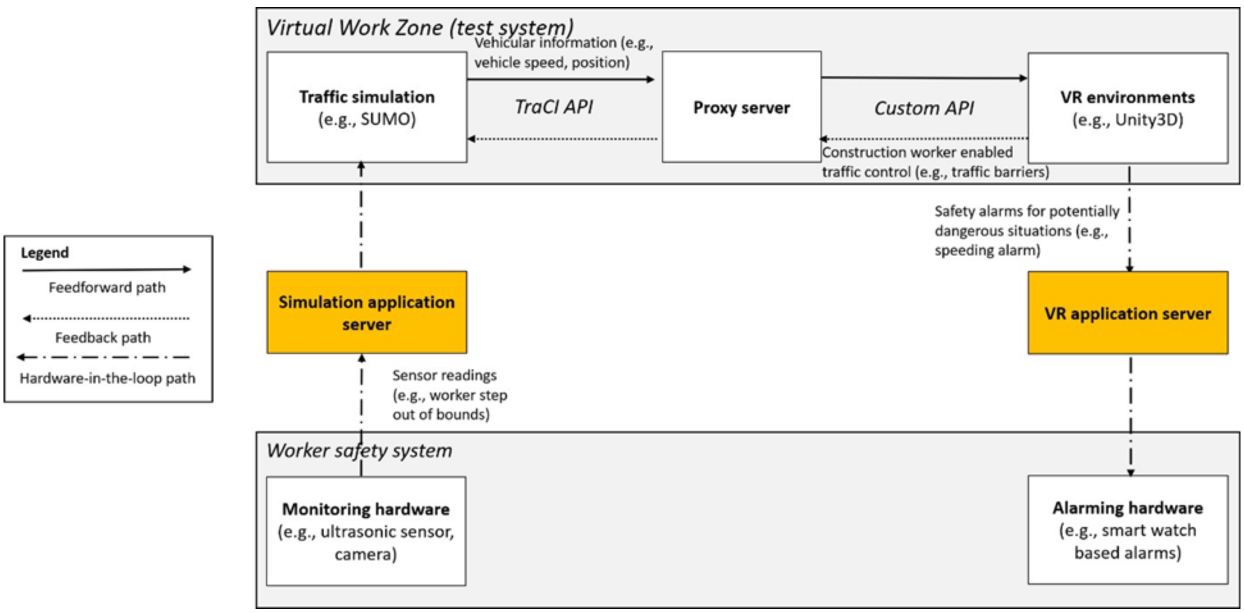


Figure 1: Hardware in the loop integration of VR, traffic simulation, worker alarm system, and work zone monitoring system

The experiment platform can also be used for HIL simulations of the smartwatch (e.g., Apple Watch) and physical monitoring hardware (e.g., ultrasonic sensors). A custom smartwatch application alerts the worker when it detects hazardous events logged by an external simulation application server. Hazardous event data could be sent to the simulation application server from the traffic simulation software, which can simulate speeding cars and collisions. Additional monitoring hardware could also send hazardous event data to the simulation application server. For example, ultrasonic sensors placed near the worker in the lab could detect when a worker had physically moved out of the safe work area and trigger smartwatch alarms. Both physical monitoring hardware and simulated traffic events can trigger smartwatch alarms, which could impact the worker's behavior in the VR environment and result in changes in simulated traffic. A worker's response to the smartwatch alarm interface (e.g., tapping the watch's screen) is also logged in the simulation application server. The suite of smartwatch alarm and other monitoring hardware make up the worker safety system being tested on the VR-HIL platform. More details can be found in previous publications (Ergan et al. 2022).

An Overview of the IRB-approved User Study and Scenarios Implemented

The research team implemented multiple traffic scenarios in VR using gaming engines and integrated them with the micro simulation platform. These VR models have been used as part the designed user studies that were conducted at NYU Building Informatics and Visualization Lab (biLAB). An overview of the designed user study is provided in Figure 2. Invited participants were asked to wear a head-mounted display to be immersed in the implemented traffic scenarios (Figure 2a), a smart watch to capture their heart rate variations and send alarm notifications (Figure 2b), a heart rate monitoring device (i.e., E4) to track their biometric data variation over time with respect to alarm periods (which could be an indication of situational awareness), and utilize a pair of controllers to navigate in and interact with the virtual environments. Smartwatch is capable of receiving heart beat data of participants throughout the experiments and sending alarms with various configurations. Each alarm a participant receives has a modality, which is defined as the stimuli an alarm uses to warn workers (i.e., vibration, vibration + sound), a frequency, which is defined as the number of repetitions an alarm is provided to a worker after a triggering alarm, and varies between 1 and 3 repeats and a duration, which is defined as the length of time the stimuli (e.g., sound +vibration) is given to a worker and ranges between 1.8 to 3.8 seconds; and a pause period, which is defined by the length of time the stimuli pauses between repeats, and ranges between 0 to 2 seconds. The integrated platform overviewed above has been deployed to maintain bidirectional data flow between the traffic simulation and the VR based on participants' actions in the VR scenarios (Figure 2c).

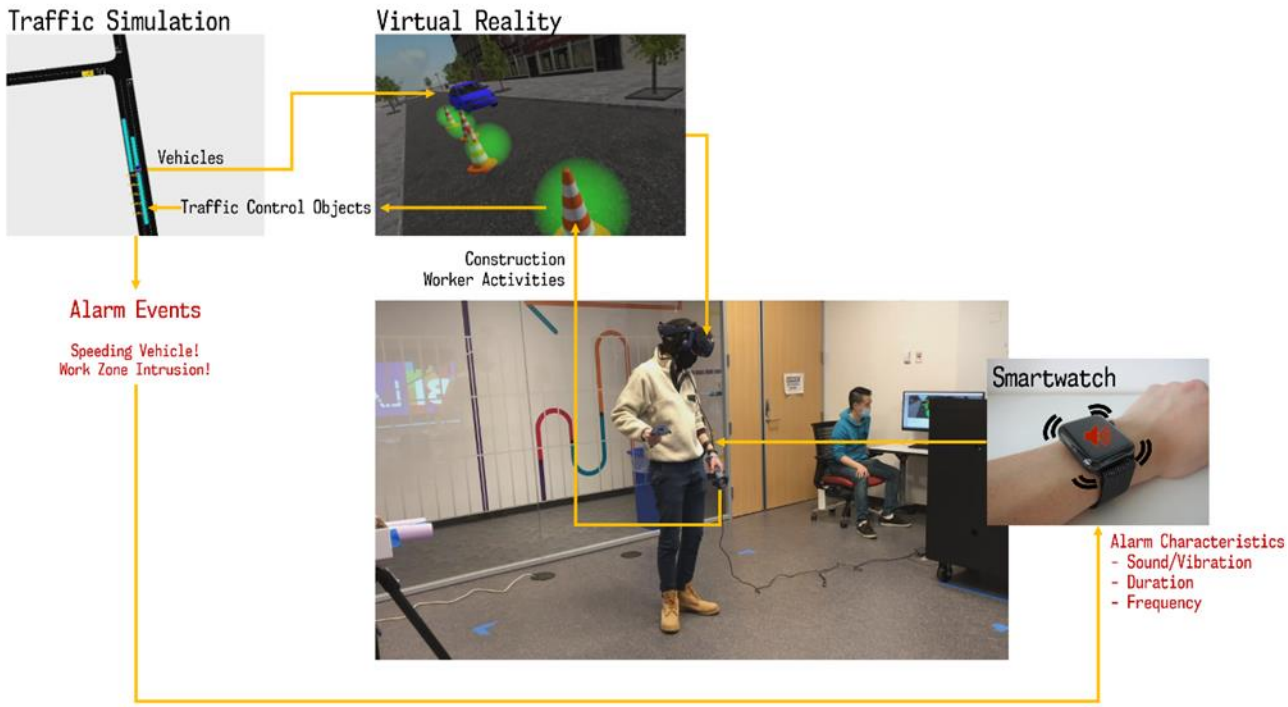


Figure 2: User study setup utilizing the integrated platform

Each participant had to perform the several tasks subject to various alarm events and smartwatch alarm characteristics depending on the scenario Three scenarios were developed in VR: Scenario 1 simulates a virtual world for an unstructured mobile work zone along an urban street, Scenario 2 is a virtual world for an urban highway work zone replicating a sensor installation task on the highway, and Scenario 3 is a virtual world for an urban intersection work zone replicating the tasks of construction surveying.. New York University's IRB approved the VR user studies (IRB-FY2020-3946) and the research team conducted them with a total of thirty-three participants for Scenario 1 ten participants for Scenarios 2 and 3. All subsequent sections discuss the data collected from the initial cohort of thirty-three participants in the first Scenario.

In Scenario 1, participants were asked to place six traffic cones in an unstructured mobile work zone along an urban street (Figure 3).

In each trial of the VR user study with Scenario 1, participants had to move a cone to each of the blinking white/yellow lights visible to participants in the virtual environment. Cones placed in the correct location would be registered by the traffic simulation, and subsequent vehicle behavior would treat the cones as obstacles to avoid. Vehicles would subsequently move down the opposite lane of the street. Each trial concluded when all six cones were placed in the correct location, as registered in the VR environment when the white/yellow lights turned green (Figure 3). Each participant conducted this sequence of work tasks twice in order to experience two modalities of alarms (vibration only, and sound with vibration combined). The order of modalities was randomized among participants.

Figure 3: A closeup view in virtual world for Scenario 1

Before the actual trials where data was recorded, participants had a chance to conduct these tasks in a training session with help from research personnel and without simulated traffic or triggered alarms in order to help them familiarize themselves with the VR environment and controllers. Only after participants stated they felt comfortable with the VR system and performing the tasks absent of traffic did the experiments proceed.

During a user study, three possible events would trigger an alarm on the smart watch.

- Trigger Event 1: Speeding Vehicles: Speeding zones in SUMO traffic simulation were set up to detect speeding vehicles (Figure 4). When a vehicle's speed exceeds 25 miles per hour (~12 m/s), a speeding vehicle alarm is immediately raised and triggers an alarm to be sent to the smart watch. In Scenario 1, three of five cars were set to be speeding vehicles and were spaced such that the speeding alarms would be raised at least 15 seconds apart.

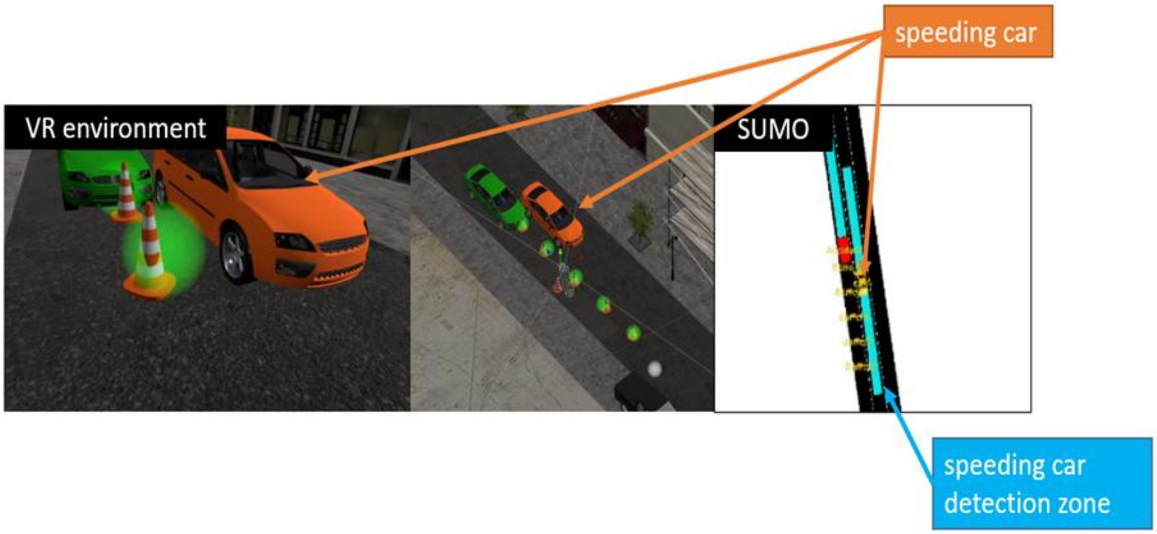


Figure 4: A closer view from the virtual world showing speeding cars and their detection zones

- Trigger Event 2: Collision Vehicles: After the first 10 seconds of the trial, a car in SUMO and VR was set up to collide with the work zone (Figure 5). An alarm was raised right when the car intruded a virtual work zone.

Figure 5: A closer view from the virtual world showing an example collision vehicle

- Trigger Event 3: Out of bounds: If a worker stands at the edge of the street's opposite lane, the ultrasonic sensor tracking the worker's physical location would trigger an alarm (Figure 6). Participants were made aware of the ultrasonic sensor before the actual trials. However, since the alarm sounds and vibrations were randomized, participants had to realize that their physical location caused the alarm completely on their own. Only three participants triggered the ultrasonic sensor's perimeter alarm. Given this very limited amount of data for Perimeter alarm events, only Speeding and Collision alarm events are analyzed in this report.

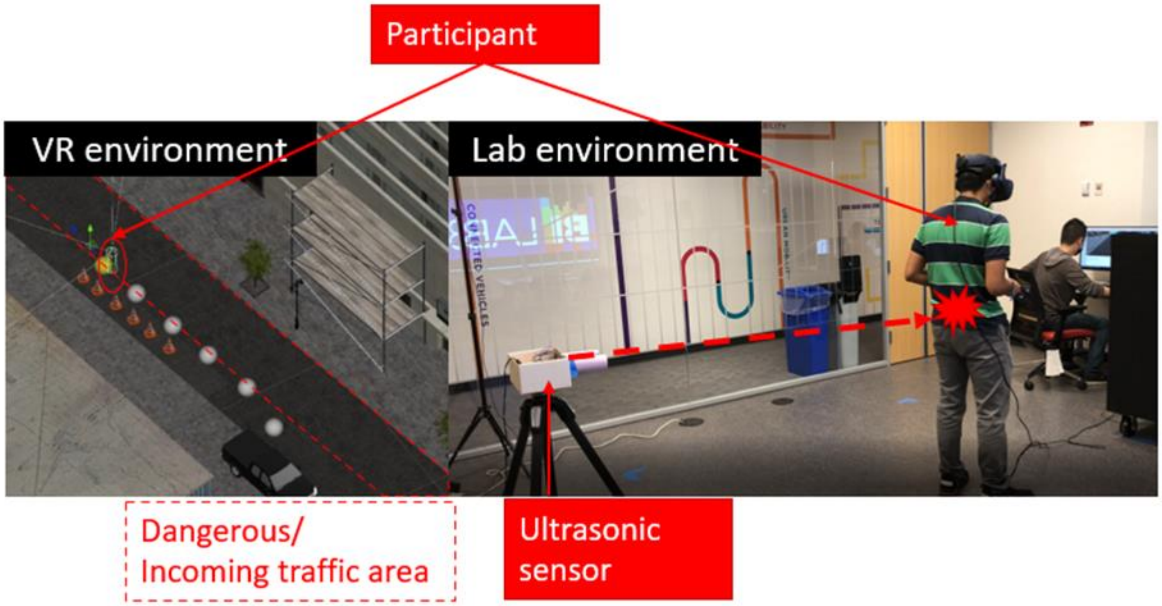


Figure 6: A photo from the real experiment boundary tracked with the ultrasonic sensors

Regardless of these three possible trigger events, the smartwatch alarm app would select a randomized sequence of alarm sounds and vibrations. In other words, the current experiment design intended there to be no easy way for participants to distinguish the trigger event cause of the alarm from the sounds and vibrations alone. Time stamps of each alarm event are recorded in the Simulation Application Server (Figure 1). As explained before, these alarm events trigger an alarm of random characteristic (e.g., duration, frequency). To emphasize, participants were told that the alarms they would receive on the watch had to be caused by one of the above three trigger events but the characteristics of the alarms was randomized and had no relation to their cause. That said, the cause of the alarm is recorded in the attention monitoring system and is analyzed in relation to how people reacted.

An overview of the data collected during the user studies

To determine the situational awareness of workers at work zones with respect to triggering events, the research team defined several features that should be considered while determining the attention range of workers. Attention is first detected when a worker acknowledges that alarm on the smart watch. However, this is not the only way to detect that workers are aware of the situation they are in. Even if they do not acknowledge the alarm by tapping the smart watch to stop the alarm, they can show other behaviors that are safe beyond the alarm acknowledgement (see Figure 7). These features have been identified as the followings:

- whether a participant acknowledged an alarm on to the smartwatch user interface
- gaze direction
- visual attention to potential hazards (e.g., cars, trucks, etc.)
- overall body position and movement
- heart rate variability (HRV)

Past research has examined the importance of these features on the construction worker safety (Luo et al. 2016; Jeelani et al. 2019; Lee et al. 2017). While the relative importance of each factor as they related to a worker's overall safety is not well understood, RL agent should at least consider data relevant to each factor to inform how a smartwatch agent can learn and adjust its alarm settings. Data on a particular worker's behavior and physiology should then be benchmarked against standard values to evaluate whether that data indicates a safe or unsafe level of worker attention. Figure 7 shows illustrations of worker behaviors under triggered events that will be considered as features for worker attention to dangerous situations.

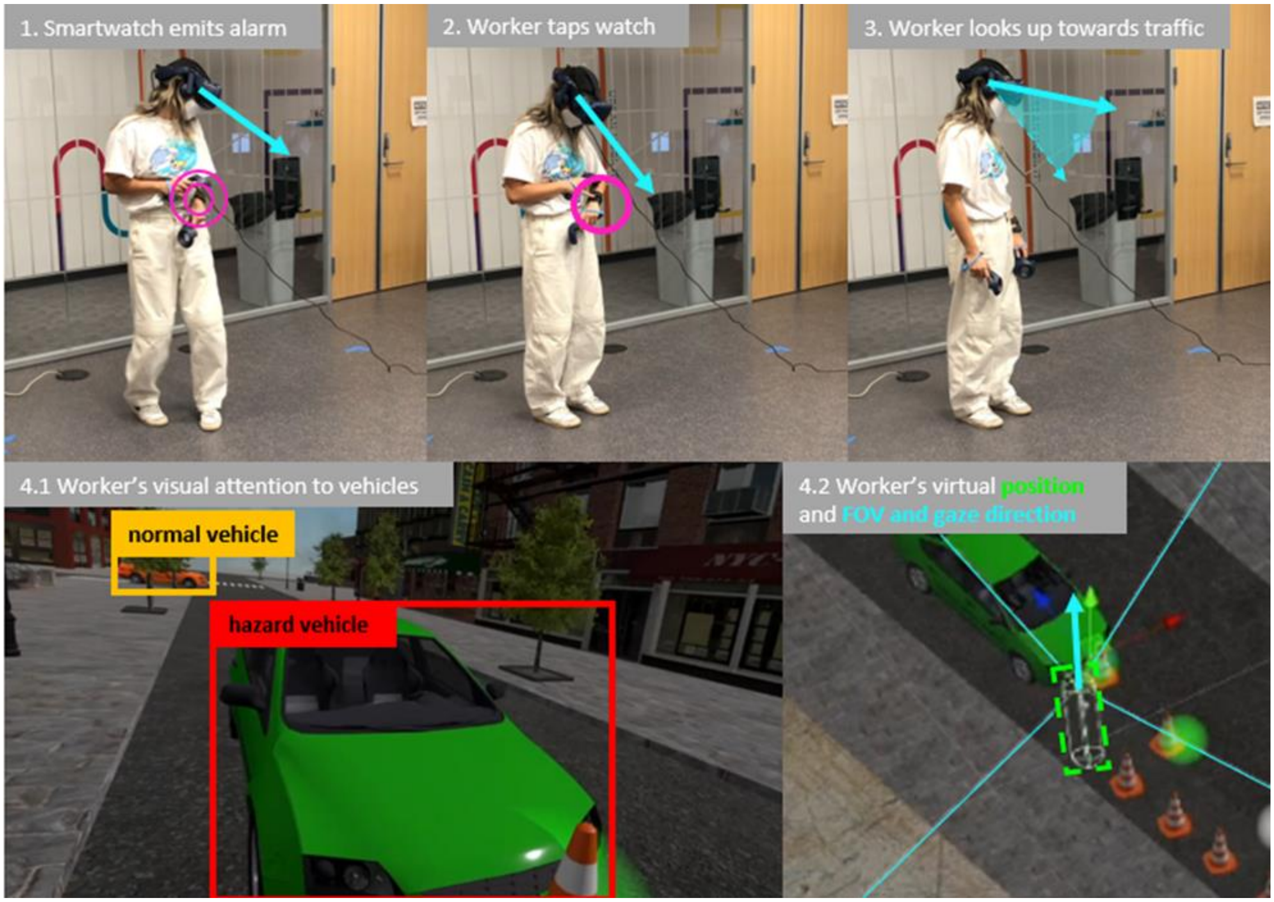


Figure 7: Illustrations of features that are included in worker attention towards safety alarms

Acknowledgement of alarms on smart watch interface: Both the alarm triggering events (i.e., speeding, collision, out of bounds) and worker responses (i.e., acknowledged, not-acknowledged) on the smartwatch interface are tracked by the smartwatch and logged in the simulation application server. If a user taps the smartwatch screen to dismiss an alarm, the response time can be calculated as the time elapsed between the time of dismissal and time when the alarm is triggered.

Gaze Direction in the VR Platform: VR platforms also allow developers to track the user's camera orientation as rotations which are quantified as quaternions by default. Each camera rotation, R_t, at time, t, can be logged by the VR platform as the rotation's quaternion, where a quaternion is a 4-dimensional representation of rotations in 3D space. Quaternion definition is a well established mathematical convention that is used by Unity and Python libraries for 3D rotations and details can be found here.

$$R_t = Quaternion(W_t, X_t, Y_t, Z_t)$$

Quaternions, however, cannot be interpreted directly as the gaze direction, g, the unit vector that represents where the worker and their VR camera is looking at (Figure 8). If the initial gaze direction at the start of the VR simulation is known, VR platforms like Unity3D have an API to convert quaternions to a rotation matrix and use that matrix to transform an initial gaze direction.

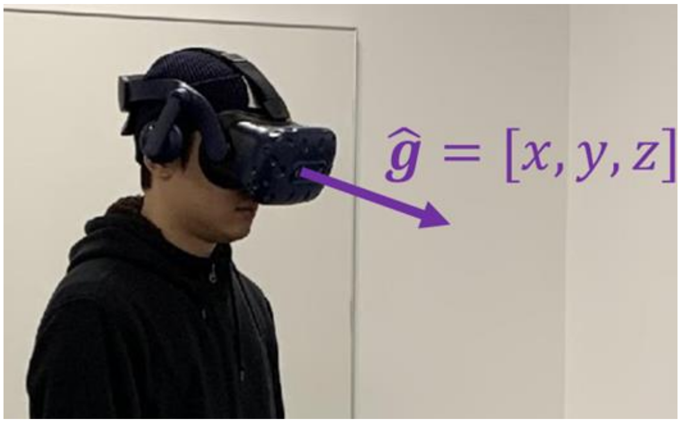


Figure 8: Gaze direction vector of the worker as tracked by the VR headset

For our current experiments, the initial gaze direction, $\hat{\boldsymbol{g}}_o$, is represented by the positive Z-axis direction:

$$\hat{\boldsymbol{g}}_o = [0,0,1]$$

Then the worker's gaze direction can be calculated by applying the rotations at time t, $\boldsymbol{R_t}$**,** to the initial gaze direction:

$$\hat{\boldsymbol{g}}_t = \boldsymbol{R_t}\ \hat{\boldsymbol{g}}_o = [x_t, y_t, z_t]$$

Gaze direction can then be timestamped and sent by the VR platform to the simulation application server. The current VR platform records the gaze direction at roughly 40Hz. The gaze directions and quaternions have been used to find if a worker was looking at the direction of the traffic (hence notices the triggering event) even if s/he did not acknowledge the received alarm on the smart watch interface. A sample data plotted for the calculated gaze directional changes over time is provided in Figure 9.

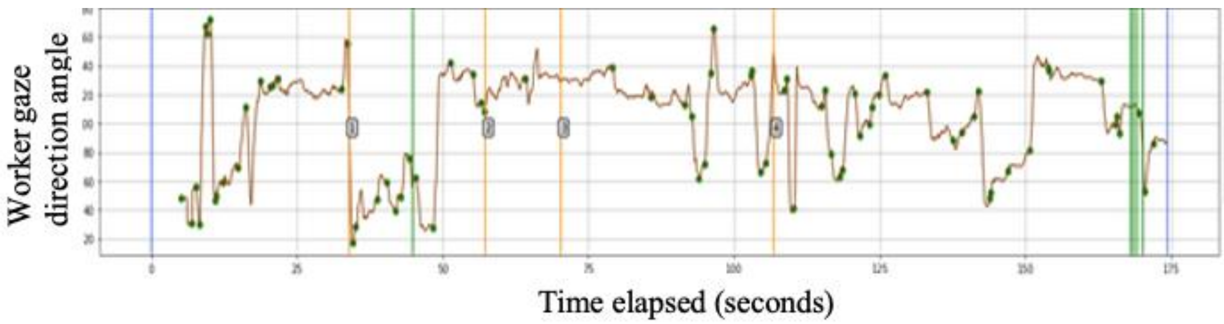


Figure 9: A sample data from a participant showing the calculated gaze direction angles. Note that the green dots on the plot are indicating head turn speed events, vertical orange lines are indicating triggered events.

__Detecting Potential Hazards in VR:__ Various VR headsets have eye tracking hardware and software capabilities. The currently available system for the research team, however, does not have any eye tracking capabilities. That said there are other methods to track the worker's area of focus and determine whether the worker paid attention to potential hazards like cars or trucks.

In the current system, the raytracing between the VR camera focal point and the worker's entire field of view is used to track whether vehicles are seen by workers. While this cannot confirm if the worker was specifically looking at a particular hazard in the general area, ray tracing detection results on the VR display, particularly in the moments after each alarm, is currently utilized to quantify a worker's attention. When we calculated this feature, we differentiated what a worker detects within their field of view as potential hazards (cars that are in their field of view) and the triggering event of speeding and collision cars. A sample data plotted for the calculated number of cars seen within the field of view of workers over time is provided in Figure 10.

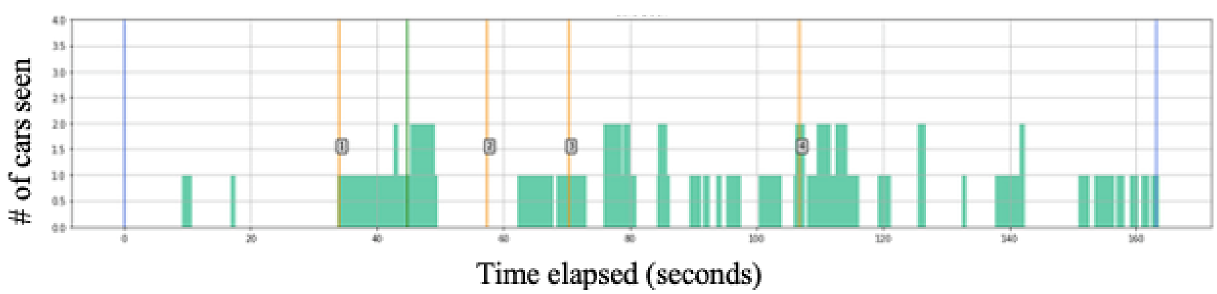


Figure 1: A sample data showing the calculated number of cars in field of view of a worker. Note that the vertical orange lines are indicating triggered events

<u>Body Position and Movement - VR Platform:</u> VR platforms such as Unity3D allow developers to continuously track a user's headset position within the virtual world as X,Y,Z coordinates. These values are world coordinates by default, meaning they can be relative to some faraway origin point in the VR environment space rather than some meaningful location in the scenario. In Unity3D, workers walk on the X-Z plane with the Y-axis as the "up/down" direction (Figure 11). These coordinates are continuously logged on the simulation application server with timestamps. The current VR platform records the worker's position at roughly 40Hz. The recorded data has been used to derive the original location and movement along the work zone over time given a triggered event timestamp.

Figure 2: Screenshot of Unity3D worker's VR perspective (left) and bird's eye view (right). Note the red and blue arrows on the right top corner, which indicate the X-axis and Z-axis location of the worker in the VR environment

A sample data plotted for the calculated worker positional changes over time is provided in Figure 12.

Figure 3: A sample data from a participant showing the calculated road positions. Note that the green dots on the plot are indicating movement events, vertical orange lines are indicating triggered events

Heart Rate Variability through the PPG Wristband: A photoplethysmography (PPG) wristband uses electrical signals derived from light reflected due to changes in blood volume pulse (BVP) during heart activity. Additional signal processing algorithms then calculate heart rate (HR) and inter beat interval (IBI) based on those BVP measurements. HR and IBI are metrics that show change of heart rate activity over time. The current experiments utilized the Empatica E4 wristband which can record all three quantities (BVP, HR, and IBI) and send that timestamped data to the simulation server. All data collected in these sensors can be displayed for each trial of each user's VR experiment session. A sample data for tracked and calculated IBI is provided for a worker in Figure 13.

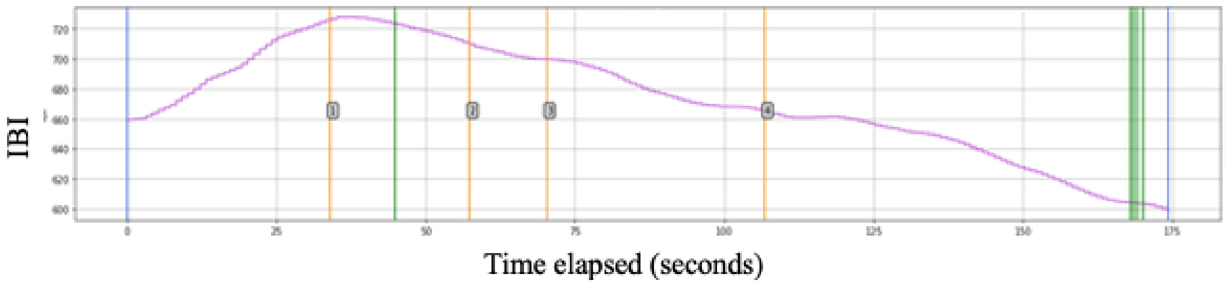


Figure 4: A sample data from a participant showing the calculated IBI. Note that the vertical orange lines are indicating triggered events

Types of Alarms Generated via the Smartwatch

Safety notifications provided to participants through the smartwatch were developed relatively independent of the specific scenarios the participants would encounter. Given the limited knowledge of worker alarm responses, the research team set out to develop and test a wide variety of alarm characteristics. The Apple Watch operating system, watchOS, has its own application programming interface (API) for emitting nine built-in types of alarms of various sounds and vibration, each lasting around 0.8 seconds. Among these types, the Retry built-in type was chosen since it had the most intense sound and vibration pattern. Previous alpha tests of the Apple Watch informed this selection as being one of the more noticeable and intense alarm types rated by watch users.

In order to test different alarm durations and frequencies, this single built-in alarm type had to be repeated multiple times with a variety of different pauses in between. Strictly speaking, the Apple Watch API does not allow any application to emit a single built-in alarm type for a completely continuous duration. The smallest "gap" in time allowed between two consecutive built-in alarms is 0.2 seconds. This study treats consecutively emitted built-in alarms as an effectively continuous alarm duration. For example, a 0.8 second duration built-in alarm emitted 3 consecutive times is considered in this study to have an alarm duration of 2.8 seconds. Figure 14 illustrates how the alarm duration is defined relative to the built-in alarms' duration.

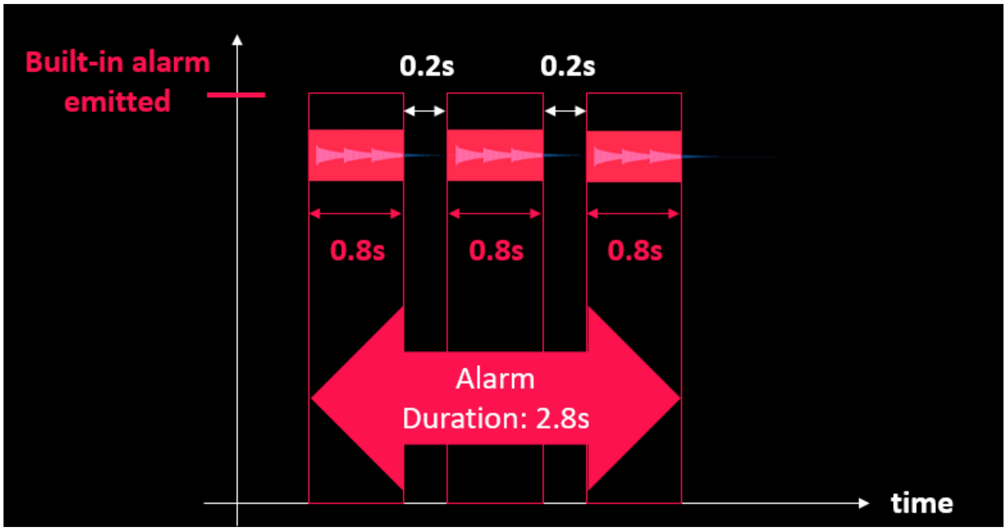


Figure 5: Alarm duration defined relative to the built-in alarms' duration

To test a variety of alarm frequencies, the Apple Watch application can emit a single *alarm duration* at multiple *repetitions* and *pause periods*. The total length of time of the *notification period* is the sum of the alarm durations and pause periods. This will be will be referred to as the *total notification period*, as illustrated in Figure 15 below.

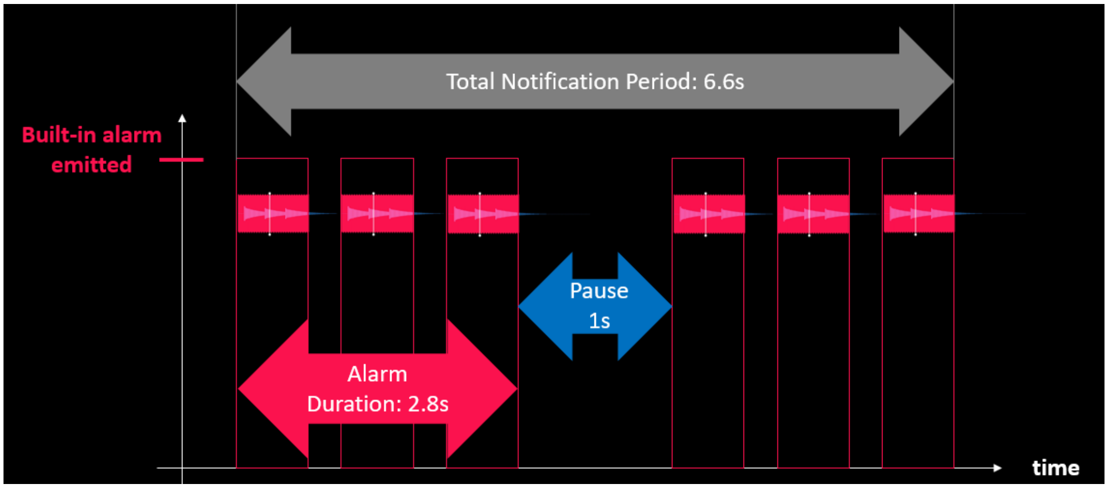


Figure 6: Example showing alarm durations, pause periods with respect to total notification period (with the number of repetitions as 2)

The Apple Watch application for this study can emit fifteen different notifications to test all different combinations of the alarm durations, pause periods, and repetitions (frequencies) given in Table 1. When an alarm event (e.g., speeding car) occurs, the Apple Watch application selects a random notification, thereby emitting a random alarm duration, pause period, and frequencies.

Table 1: Notifications (modes) implemented in user studies

Notification # (Mode)	Alarm Duration (s)	Pause Period (s)	# of repetitions (Frequency)	Notification Period (s)
	D	*P*	*R*	*R*D+(R-1)*P*
1	1.8	0	1	1.8
2	2.8	0	1	2.8
3	3.8	0	1	3.8
4	1.8	1	2	4.6
5	2.8	1	2	6.6
6	3.8	1	2	8.6
7	1.8	2	2	5.6
8	2.8	2	2	7.6
9	3.8	2	2	9.6
10	1.8	1	3	7.4
11	2.8	1	3	10.4

12	3.8	1	3	13.4
13	1.8	2	3	9.4
14	2.8	2	3	12.4
15	3.8	2	3	15.4

While the alarm durations for this study were selected based on previous alpha tests with the Apple Watch in isolation, the frequencies and pause periods were selected to test if these alarm variations could produce wider variety of worker reactions. The notifications in this current Apple Watch application do not account for any expected time frame of human responses to alarms. As explained earlier, the triggering events (i.e., speeding car, collision car, out of bounds) did not predetermine the notification type emitted by the Apple Watch.

In terms of alarm modality, the Apple Watch API's alarm types can only be emitted with sounds and vibrations or with only vibrations. Switching between these modalities can only be done by a user switching on and off the Apple Watch's Silent Mode in its system-wide settings. This prevents any Apple Watch application from changing notification modality automatically and inhibited the research team from testing worker responses to different alarm modalities during a single VR simulation trial. Therefore, participants in the VR user studies conducted two trials, one for each alarm modality (sound with vibration alarms, or vibration only alarms) fixed during a single trial.

To clarify the terminology, the list of parameters that play role in defining an alarm is defined in the following short glossary:

- Modality: the stimuli an alarm uses to warn workers (i.e., vibration, vibration + sound).
- Alarm duration (D): the length of time the stimuli (e.g., sound +vibration) is given to a worker and ranges between 1.8 to 3.8 seconds.
- Frequency (# of repetitions, R): the number of repetitions an alarm is provided to a worker after a triggering alarm, varying between 1 and 3 repeats.
- Pause period (P): the length of time the stimuli pauses between repeats, and ranges between 0 to 2 seconds.
- Notification/mode and its period: Notification (also referred to as mode) is a combination of alarm durations, frequencies, and pause periods. There are 15 modes tested in this study. Notification period is the total duration elapsed between the first alarm emitted after a triggering event and the last alarm duration lapsed including the pause periods in between, and is calculated as $R*D+(R-1)*P$.

Initial Statistical Analyses on Alarm Characteristics

Preliminary analysis evaluated the variation in features defined earlier in the report for detecting attention and situational awareness of workers across alarm characteristics, assuming they were independent variables:

- Alarm modalities: Sound and vibration combined, or vibration only. Given that there are only two sample groups for analysis, a one-way analysis of variance (ANOVA) F-test was used to determine if the sample distributions were significantly different.
- Alarm duration: 1.8, 2.8, or 3.8 seconds. Given that there are more than two sample groups for analysis, a Kruskall-Wallis H test was used to determine if the sample distributions were significantly different.
- Alarm frequencies: 1, 2, or 3. Given that there are more than two sample groups for analysis, a Kruskall-Wallis H test was used to determine if the sample distributions were significantly different.
- Alarm pause periods: None when emitted once, 1 second or 2 seconds when repeated twice or three times. Given that there are more than two sample groups for analysis, a Kruskall-Wallis H test was used to determine if the sample distributions were significantly different.
- Number of previous alarms: Number of previous alarms experienced by the worker beforehand within the VR trial. This variable was analyzed specifically because alarm fatigue is a notable issue with safety alarms. While certain users ended up experiencing 9 alarms, only sample groups of users who experienced 0 to 5 previous alarms were included in this analysis, since Kruskall-Wallis requires sample sizes be at least 5 or greater.

Boundary Set for the Reinforcement Learning Model

Reinforcement learning (RL) is a machine learning approach where an agent can learn to maximize a reward by observing and interacting with an environment without prior external input guidance (i.e., not supervised learning). In the context of this project, the agent is a context aware and intelligent version of the alarm application used on the smartwatch in previously conducted VR user studies. Instead of emitting random alarm characteristics (durations, frequencies, and sounds/vibrations), the alarm agent will choose a set of alarm characteristics to emit, referred in RL as actions. These agent's actions are based on a policy, a function that predicts which alarm characteristics are likely to trigger the worker to physically react. Data calculated for the identified features of attention that RL agent considers are marked in Figure 16. These predictions of worker reactions would be based on prior data from the environment, which includes the worker wearing the smartwatch, their progress in a sequence of construction activities, and the traffic vehicles moving around the work zone. The alarm agent can store constantly provided data from the environment as observations during a learning episode.

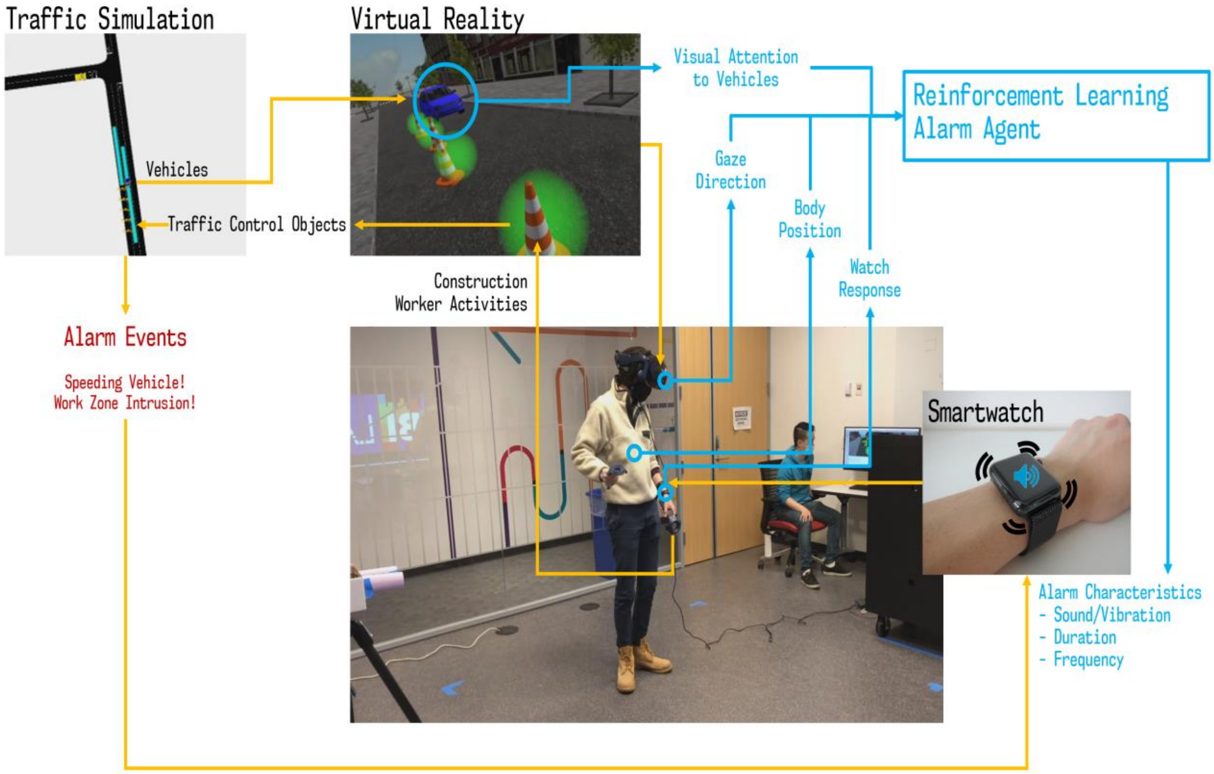


Figure 7: Illustrative showing the features of attention that RL agent considers

More importantly, the alarm agent will focus on how workers react to the alarms to quantify an evaluation score for how effective the agent's choice in alarm characteristics were in triggering a reaction. This evaluation score derived from worker reactions serves as a reward for the alarm agent see Figure 17 for list of actions, observation states, and reward function elements). By recording these observations and rewards from the environment, the alarm agent can regularly update its predictions of worker reactions to future alarms to inform how it picks alarm characteristics to maximize reward in future episodes. Ultimately, the rewards and observations reinforce how the alarm agent optimizes alarm characteristics to promote safe worker behaviors in roadway construction sites as it observes more and more episodes of worker and work zone environment data. Therefore, the parameterization of observations and rewards will need to be carefully defined to ensure the alarm agent contributes to worker life safety in roadway construction sites. This section serves primarily as a preliminary definition of the reinforcement learning model's parameters before RL training simulations are conducted.

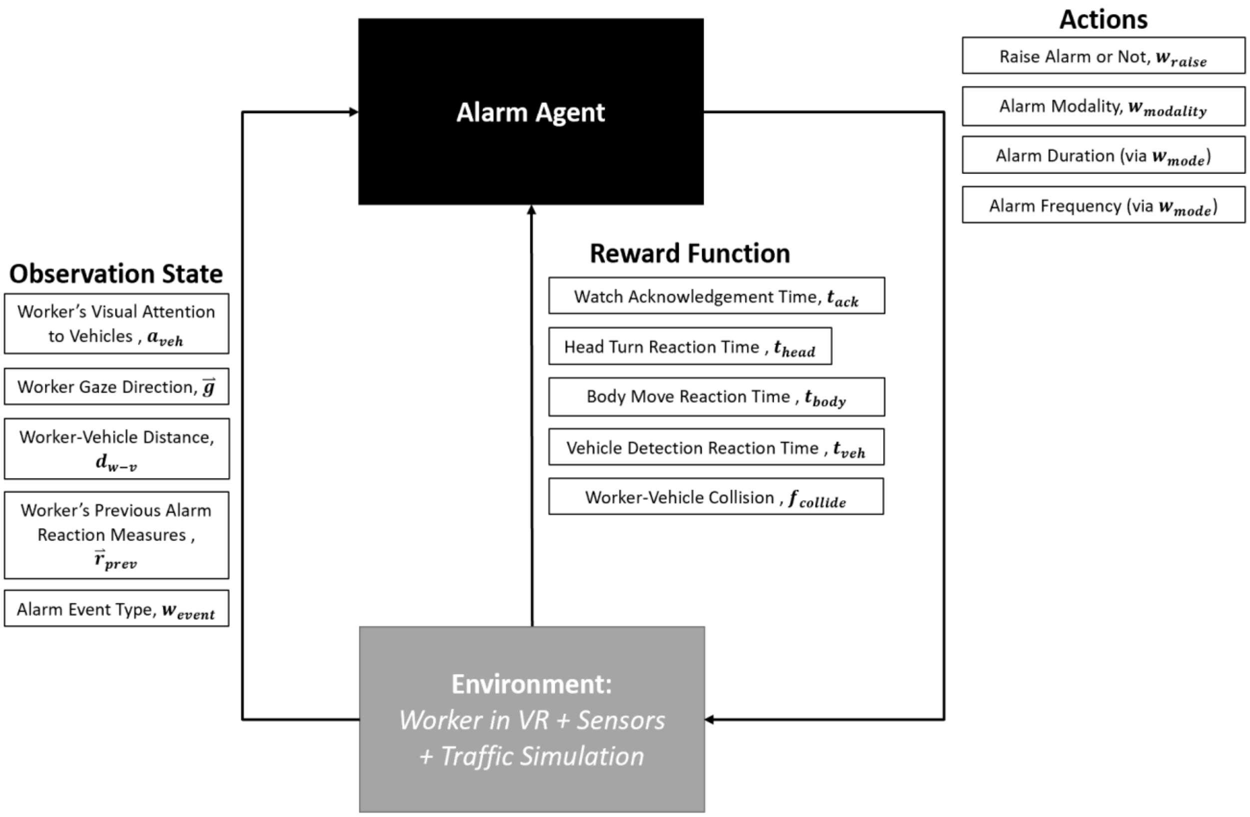


Figure 8: RL parameters defined: actions, observation states, and reward function elements

Actions defined for the RL agent:

The agent will make several decisions regarding each alarm.

1. Decide whether to raise the alarm or not, even if the traffic simulation detects a speeding vehicle, work zone intrusion, or out of bounds.

$$w_{raise} = \begin{cases} \mathbf{1}, raise\ alarm \\ \mathbf{0}, do\ not\ raise\ alarm \end{cases}$$

2. Choose an alarm modality, $w_{modality}$

$$w_{modality} = \begin{cases} \mathbf{1}, vibration\ only\ alarm \\ \mathbf{2}, sound\ and\ vibration\ alarm \end{cases}$$

3. Choose an alarm mode, w_{mode}, a value that represents a specific combination of duration, frequency, and pause period specified in Table 1 above.

The agent's actions regarding alarms, A_i, can therefore be represented as:

$$A_i = w_{raise}\ [w_{modality}, w_{mode}]$$

such that $A_i = [0, 0]$ means the alarm agent decided not to raise an alarm. All other non-zero combinations denote the agent decides to emit an alarm.

Reward function defined for the RL model:

If we use an analogy here to say that the RL agent will act like a "safety coach", reward function helps to review the different **forms of worker reactions** to alarms that the coach should *reward* workers to do more often:

1. Worker **avoids a collision** with vehicle.
2. Worker taps the smartwatch screen to **acknowledge** the alarm.
3. Worker **turns their head** (i.e., changes their general gaze direction towards the incoming traffic).
4. Worker sees cars and **detects the hazardous vehicle** (i.e., worker's field of view includes cars and the **event triggering car**).
5. Worker **moves their body** quickly away from the incoming traffic.

As the numbering above suggests, the research team is presuming a ranking for these forms of worker reactions in descending importance for their safety. This ranking will inform the overall logic for how the alarm agent calculates rewards for individual alarms, as explained later in this section. That said, it is unknown within the research literature and occupational safety databases in how some of these specific forms of worker reaction ultimately correlate to incidence of worker injury and fatality. Such correlations could be drawn after running a significant number of VR user studies where collisions could occur on the VR platform. The above ranking could then be informed by empirical data rather than the team's intuition on safe behaviors.

With the exception of worker collision with vehicles, these forms of worker reactions have been quantitatively measured using the following *time-based* metrics that evaluate the data captured in VR user studies. When each alarm is emitted, the RL agent will record experiment data as part of a **reaction history** within a certain timeframe, $\boldsymbol{t_{frame}}$, afterwards. Within this reaction history, the RL agent will calculate time-based metrics as described below. Certain metrics will involve calculations between timesteps ($\Delta \boldsymbol{t}$) of the reaction history, which is 0.02 second (50Hz) for the current VR- traffic simulation platform. Then, the RL agent will compare these metrics to an expected reaction time, $\boldsymbol{t_{exp}}$, a value that should be less than the reaction history timeframe, $\boldsymbol{t_{frame}}$.

The following metrics has been defined in this study to be used by the RL agent for each form of a worker's alarm reaction.

1. Worker avoids collision with vehicle

   Collision Boolean, $\boldsymbol{f_{collide}}$ - In the VR engine, Unity, the worker's entire body is idealized as a 3D capsule geometry and set as a "Collider" object. Vehicles have 3D Box geometry colliders as well. Unity constantly calculates and checks when any of these Collider geometries *intersect* and can then determine the boolean value, $\boldsymbol{f_{collide}}$**,** representing whether a worker's body collides with a vehicle at any moment.

   $$\boldsymbol{f_{collide}} = \mathbf{1}\ if\ worker\ collides\ with\ vehicle, \mathbf{0}\ otherwise$$

   Of note, unlike the subsequent values below, this parameter is always updated, regardless of whether an alarm is emitted or not.

2. Worker taps the smartwatch screen to acknowledge the alarm

   Acknowledgement Time, $\boldsymbol{t_{ack}}$ - The time elapsed between when the alarm was emitted on the watch and when the worker tapped the watch screen to acknowledge the alarm. An example calculation of this metric is illustrated in Figure 18. If the worker ignores the alarm, then $\boldsymbol{t_{ack}} = \boldsymbol{t_{frame}} + 1$, which will exceed the expected reaction time $\boldsymbol{t_{exp}}$**.**

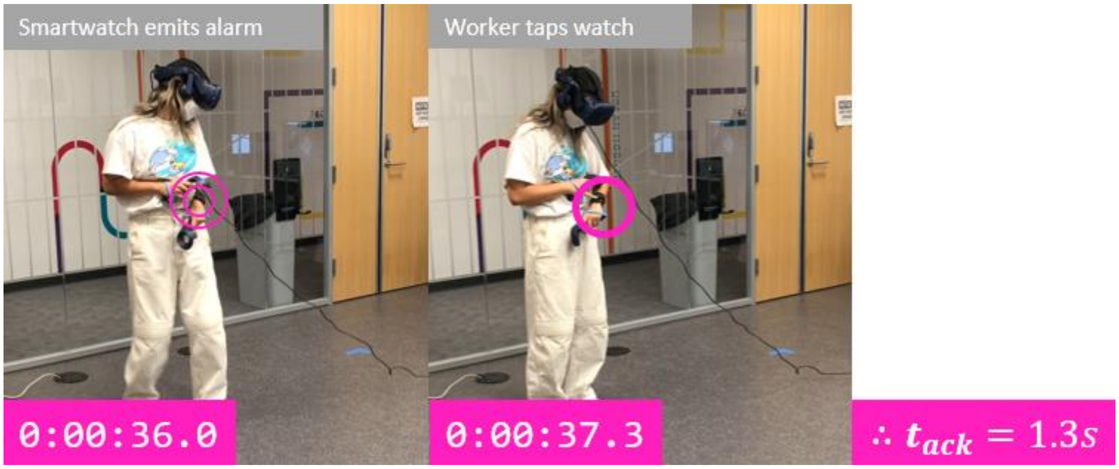


Figure 9: Illustration of how watch response/acknowledgement time is calculated

3. Worker turns their head (i.e., changes their general gaze direction)

Head Turn Reaction Time, t_{head} - The time elapsed between when the alarm was emitted on the watch and when the worker performed the *maximum change* in gaze direction *between equal timesteps* and *within a timeframe* t_{frame} after the alarm. Note that t_{head} will generally be less than timeframe t_{frame} but not necessarily less than the expected reaction time t_{exp}. Figure 19 shows an example calculation of this metric, in blue, assuming $\Delta t = 0.5$ second timesteps and $t_{frame} = 1$ second timeframe. Actual calculations will be done between 0.02 second timesteps.

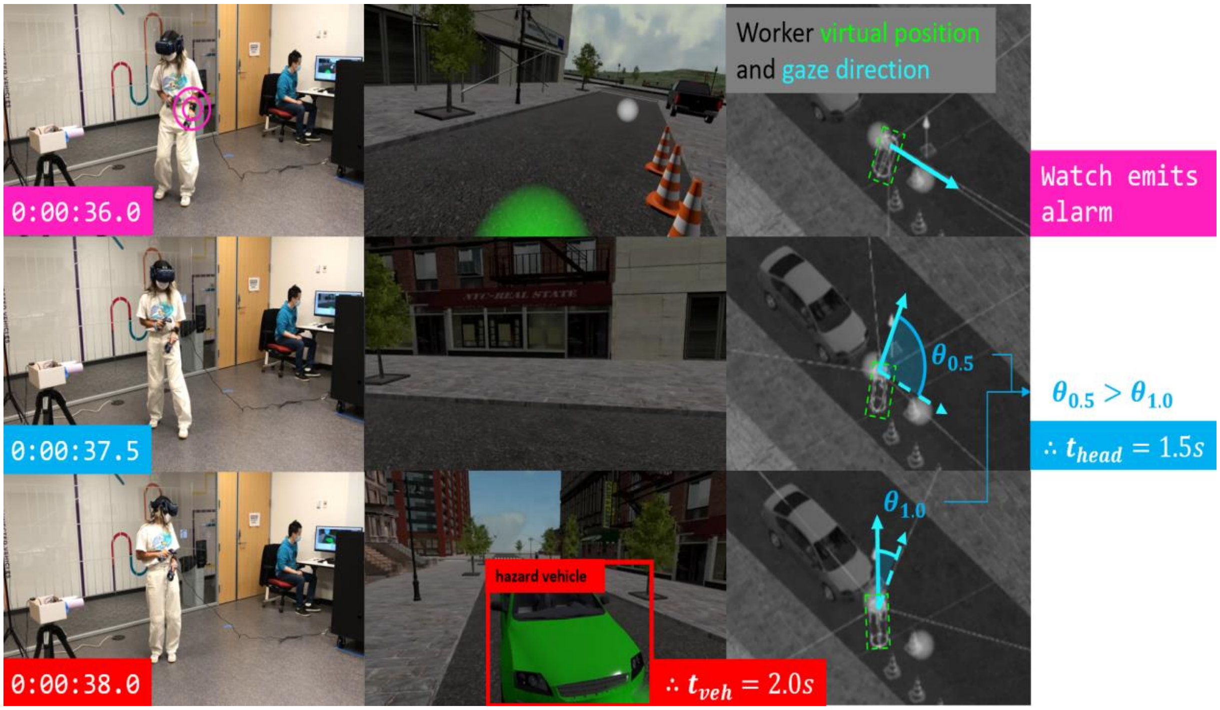


Figure 10: Illustration of how Vehicle Detection Reaction Time, t_{veh}, and Head Turn Reaction Time, t_{head}, are calculated

4. Worker sees the hazardous vehicle (i.e., worker's visual attention detects a vehicle).

Vehicle Detection Reaction Time, t_{veh} - The time elapsed between when the alarm was emitted on the watch and when worker *first* sees the hazardous vehicle in their field of view. Figure 19 shows an example calculation of this metric, in red. If the worker never sees the hazardous vehicle, then $t_{veh} = t_{frame} + 1$, which will exceed the expected reaction time t_{exp}.

5. Worker moves their body quickly.

Body Movement Reaction Time, t_{body} - The time elapsed between when the alarm was emitted on the watch and when the worker performed the maximum change in his/her position in the virtual work zone *between equal timesteps*. Note that t_{body} will generally be less than timeframe

t_{frame} but not necessarily less than the expected reaction time t_{exp}. Figure 20 shows a hypothetical example calculation of this metric, assuming $\Delta t = 0.5$ second timesteps and $t_{frame} = 1$ second timeframe. Actual calculations will be done between 0.02 second timesteps.

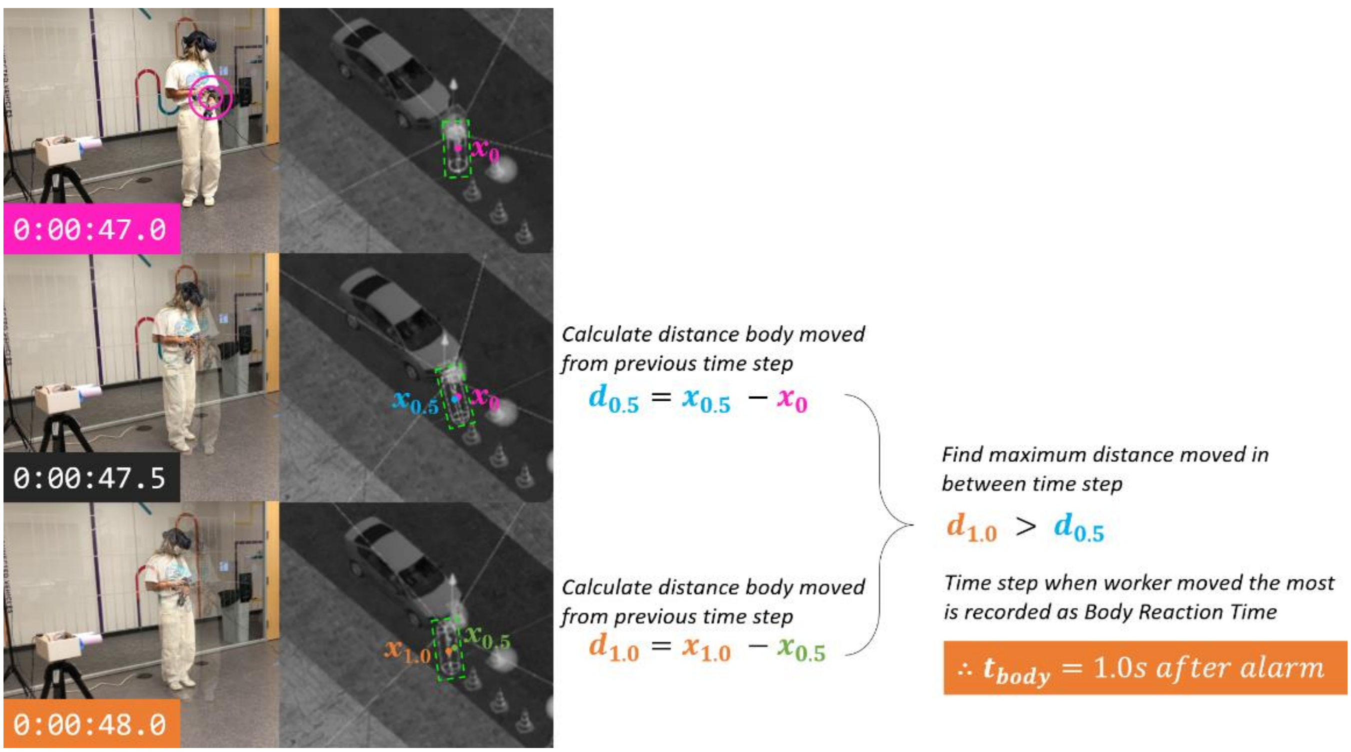


Figure 11: Illustration of how Body Movement Reaction Time, t_{body}, is calculated

Conventional RL models formulate the reward function to calculate a single scalar value for an agent to maximize. As this project is comparing multiple dimensions of a worker's reaction to an expected reaction time, there is still the matter of using the separate comparisons to formulate a single reward value, R_i, calculated after each alarm. Numerous approaches have been suggested for this general problem of multi-objective reinforcement learning. For now, this project will consider a cascading hierarchical decision tree structure for calculating a single scalar reward, as illustrated in Figure 21.

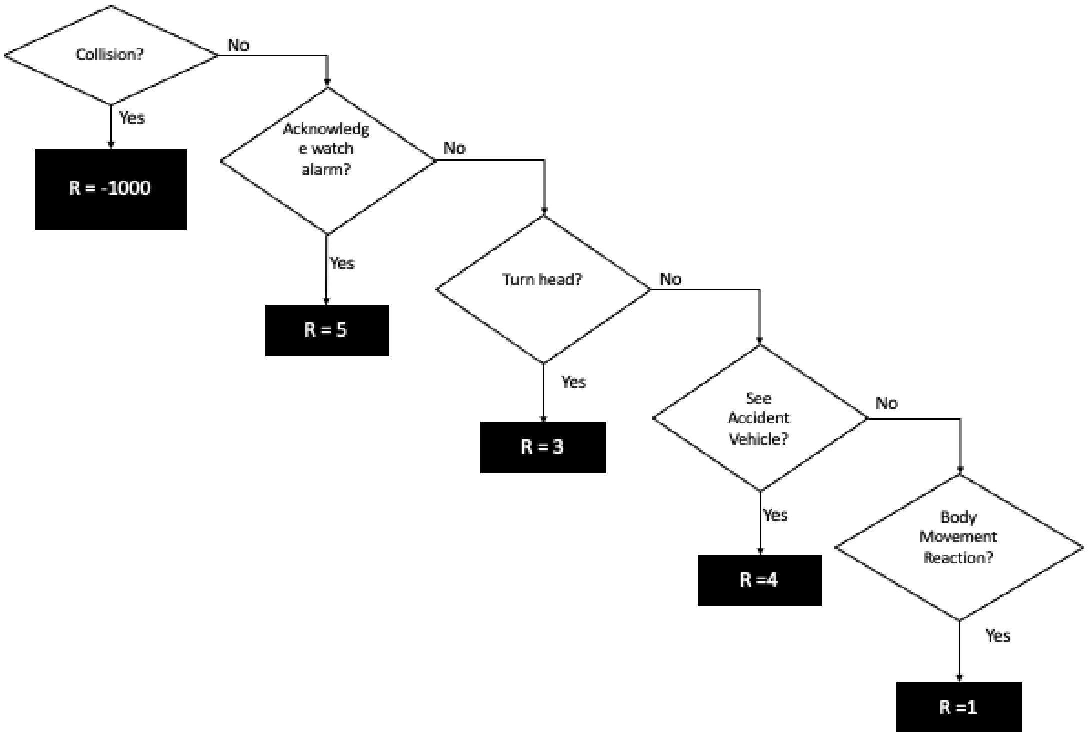


Figure 12: Reward function hierarchy

The hierarchy is based on the presumed ranking presented at the beginning of this section. If the worker collides with the car, a large negative reward or cost, $-C_{collide}$, is assigned after the alarm. If not, the reaction time metrics are sequentially compared to the expected reaction time. A scalar reward value is calculated based on which reaction time metric first falls under the expected reaction time in the hierarchy. This is why body movement is last in the hierarchy. As long as the worker did not collide with a vehicle, this decision tree only rewards a worker for moving their body if they did not acknowledge the alarm on the watch, nor saw the correct vehicle, nor even turn their head. If no reaction is observed at all after the alarm, then a small negative reward, $-\varepsilon$, is the final reward resulting from that alarm. This will slightly penalize workers who do not react at all to the alarms. Preliminary cost and reward values are provided in Table 2.

Observation States defined for the RL agent:

Returning to the "safety coach" analogy again, we now identify what aspects of a worker's current behavior or condition, at any given moment, should be worth focusing on to help pick an alarm action. The observation state is a 1D vector that the RL agent will use to capture these aspects. While traffic worker safety literature does not often offer guidance to workers on specific physical movements to the level of detail familiar to professional sports athletes, safety training and guidelines do emphasize the general importance of *situational awareness* and *maintaining a safe distance* from traffic. Therefore, the alarm agent's observation state should evaluate the following:

- *How far away is the worker from the nearest vehicle?*
- *Does the worker see the correct hazardous vehicle?*
- *Does the worker at least see nearby vehicles, regardless of whether they are hazardous or not?*

Table 2: Preliminary reward and cost values in RL model training

	Reward	Cost
$-C_{collide}$		-1000
R_{ack}	5	
R_{veh}	4	
R_{head}	3	
R_{body}	1	
$-\varepsilon$		-0.1

The RL agent will evaluate these questions by analyzing worker behavior data within a fixed timeframe, $\boldsymbol{t_{frame}}$, over *previous* simulation timesteps. For now, the research team will assume this fixed timeframe is equal to the reaction history timeframe described earlier ($\boldsymbol{t_{frame}}$). For each question above, a quantity will be calculated:

- **How far away is the worker from the nearest vehicle?**
  This is evaluated by measuring and finding the minimum distance, $\boldsymbol{d_{min}}$, between worker and vehicles in the last $\boldsymbol{t_{frame}}$ seconds, thereby quantifying how well the worker has maintained distance with traffic.

- **Does the worker see the correct hazardous vehicle?**
  This is evaluated by a boolean variable, where:

$$\boldsymbol{f_{haz}} = \mathbf{1} \text{ OR } \mathbf{0}$$

  depending on whether the worker saw the correct hazardous vehicle or not in the last $\boldsymbol{t_{frame}}$ seconds.

- **Does the worker at least see nearby vehicles, regardless of whether they are hazardous or not?**
  This is evaluated by a boolean variable, where:

$$\boldsymbol{f_{veh}} = \mathbf{1} \text{ OR } \mathbf{0}$$

  depending on whether the worker saw any vehicle or not in the last $\boldsymbol{t_{frame}}$ seconds.

Additionally, the alarm agent's observation state will contain the cause of the alarm when traffic simulation detects speeding vehicles and collisions. The overall observation state can be parameterized as a vector including time elapsed in the episode and all of the aforementioned values:

$$[t, d_{min}, f_{haz}, f_{veh}, w_{event}]$$

While it is also possible for an agent to record and store a time series of observation states, an *observation state history* in RL terminology, the research team's current implementation of the RL agent only calculates and uses the most recent observation state for selecting the alarm.

Synthetic data generation:

Ideally, we could train the RL agent by running additional trials of VR user studies. Each VR user trial would be a training episode. Pairs of observation states and actions taken by the alarm agent as well as the cumulative reward at the end of the episode are collectively recorded in an **experience history.** At the end of each episode, the agent updates its policy based on this experience history to use in the next episode. However, setting up and running VR user studies to train the alarm agent can be a very time and labor-intensive activity. In the past year, the research team has managed to conduct VR user studies with 50 unique participants, far below any statistically significant sample size for adequate model training. In this section, we detail and approach we implemented to generate synthetic training data for the alarm agent. As before, the synthetic data training will still run the same traffic simulations as done in the real VR user studies. But synthetic worker behavior data has been generated by a separate machine learning model (Figure 22). Training a machine learning model to generate realistic worker behavior data is a challenge all of its own. Various models have been considered by the research team (e.g., autoencoders, transformer-attention mechanisms). Details of these models are beyond the scope of this report but an overview of how these models are trained separately and work with the RL alarm agent training loop will be described. Overall, the synthetic data generator serves as a substitute for real workers in VR for training the alarm agent.

Training and Testing:

We implemented the following simulation to train the RL agent (Figure 23).

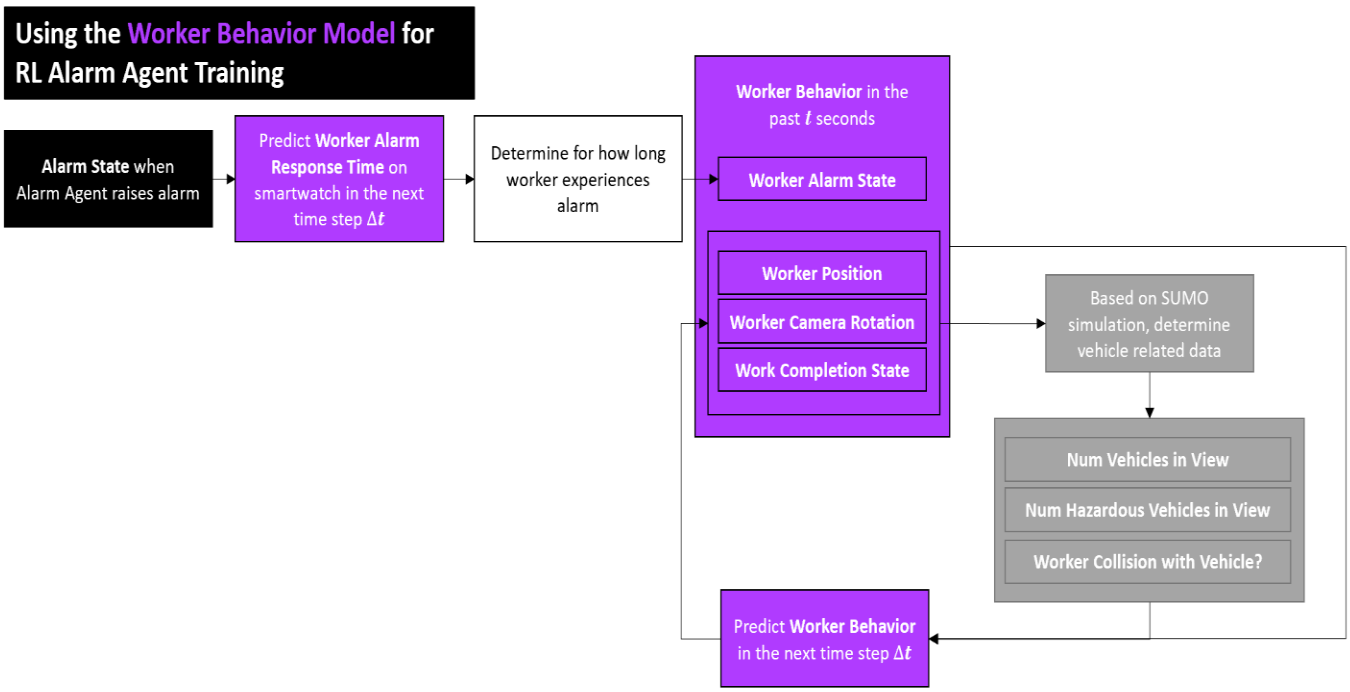


Figure 13: Prediction of worker behavior for generation of synthetic data for observation states

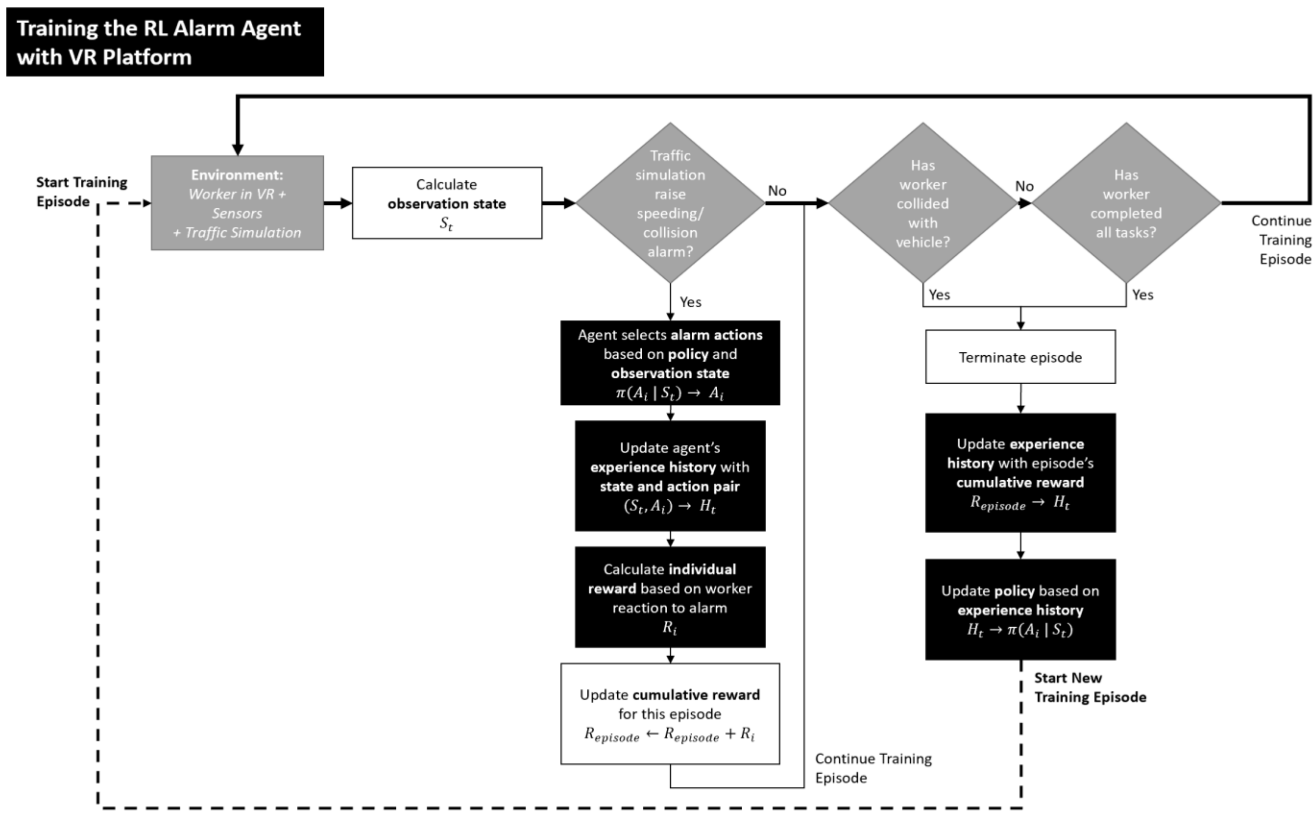


Figure 14: RL alarm agent training loop

Testing or evaluating the RL agent is the same loop, but without any updates to the policy. Typically, the average of all episodes' cumulative rewards is taken as the performance metric in for RL agent evaluation. Results of the statistical tests conducted with the data and the RL training process are provided in the next section.

Results

Results are discussed in two sections. First section provides the statistical analysis done on the captured data on VR studies. The second section provides the RL training outcomes.

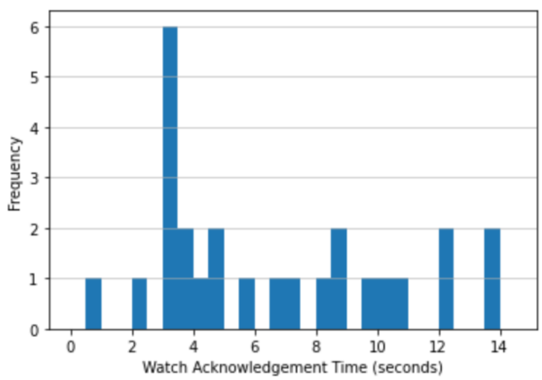
Worker Reaction Benchmark Results

We analyzed several metrics to understand the impact of changes on alarm configurations on workers' response types and durations. We defined four response types that are considered an action that is considered as situational awareness. These include (a) workers tapping the smartwatch to acknowledge received alarms, (b) workers turning their head towards the approaching traffic after receiving alarms,

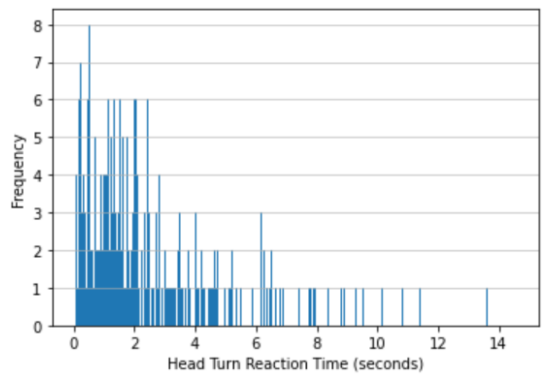
(c) workers detecting cars in their field of view after receiving alarms (and further analyzed if the cars they saw were the cars that triggered alarms), and (d) workers moving in the negative direction with respect to the normal direction of the traffic flow direction (i.e., moving away from the traffic). We analyzed the data with respect to these four actions across participants and across alarm modalities, frequencies, and durations. This section provides the results of histograms on types of actions taken over time by all participants across modalities and the reaction times for each action per alarm modality.

Analysis of response times per alarm configuration:

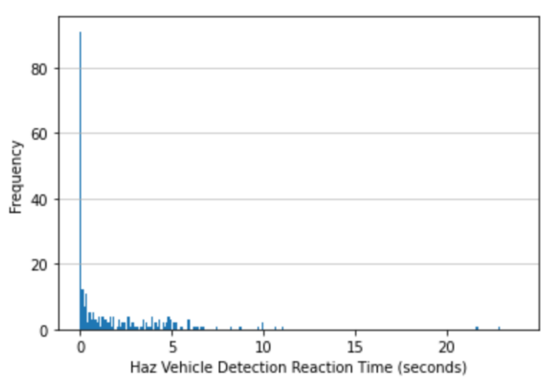
Figures 24 shows histograms of reaction times for each measure of worker reaction to all alarm configurations. Vehicle visual attention times are often close to zero because participants usually saw vehicles in their peripheral field of view. The attention monitoring system currently does not distinguish between whether the participant sees a vehicle in the center of their VR field of view or periphery.



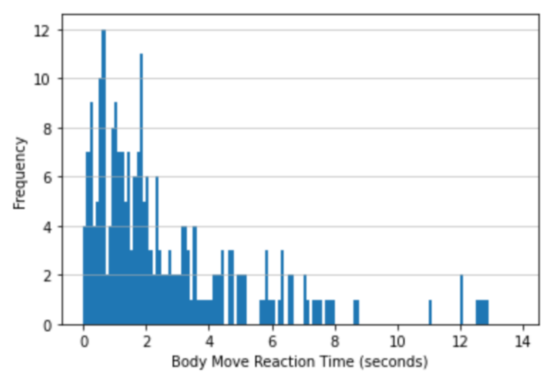
(a)



(b)



(c)



(d)

Figure 15: Distribution of reaction times across four response types for all alarm types

Figure 24 shows that reaction times across response types confine within 1.5-3 seconds timeframe in general except the watch acknowledgement. Initial analysis showed that response times do not change drastically across alarm durations and frequencies, so the plots only show response times across response types across both modalities. Table 3 shows the average and standard deviation of reaction times across worker response types. The mean reaction times in Table 3 were used as expected reaction times (E[t]) for the reward function. In the table, 'alarm acknowledged' means workers tap the smart watch to acknowledge the alarms, 'positive head turn' means workers turn their heads towards the traffic direction, 'detected vehicles in FOV' means workers see vehicles in their field of view (FOV) when they turn their heads, and 'positive body movement' means workers move away from the traffic lane. Calculations of these response times are detailed in the research method section above.

Table 3: Statistics for worker reaction times for each response type across all alarms

	Reaction Time (s)	
Response type	mean	std
Alarm Acknowledged	6.63	3.75
Positive Head Turn	2.51	2.36
Detected Vehicles in FOV	1.83	2.98
Positive Body Movement	2.7	2.6

The mean and standard deviation of reaction times for each response type (Table 3) show that workers have situational awareness before they acknowledge alarms on the watch itself. It takes less than 2.5 seconds to be aware of a problem around themselves by turning their heads, seeing approaching traffic, and changing their location until they acknowledge them on the watch. This is a supporting evidence that there are other forms of safe behaviors that can be considered to eliminate alarm fatigue at work zones before triggering alarms by studying worker reactions beyond acknowledgements.

The research team also analyzed the mean response times across response types (i.e., body movement, head turn, visual attention to vehicles, and smartwatch acknowledgement) for every alarm configuration participants received (where alarms showed variations with respect to modality, frequency, and duration- randomly). Tables 4 thru 7 show these statistics for each response type with green and red cells respectively indicating faster and slower reaction times for specific combinations of alarm characteristics.

Watch acknowledgment response times are shown in Table 4. Most apparent is the fact that a number of people did not acknowledge a number of alarm combinations. Acknowledgement times for specific alarm characteristics are not further analyzed given the lack of data on watch taps.

Table 4: Statistics for watch acknowledgement times in response to alarm configurations (modality, duration, frequency, repetitions)

modality	mode	duration (s)	pause period (s)	# of repeats	n	mean (s)	std (s)
sound + vibration	10	1.8	1	3	3	5.729	3.686
	7	1.8	2	2	1	3.292	0
	13	1.8	2	3	1	0.973	0
	1	1.8	n/a	1	1	3.001	0
	5	2.8	1	2	2	4.986	0.649
	14	2.8	2	3	4	7.281	4.387
	6	3.8	1	2	2	11.003	1.274
	12	3.8	1	3	1	3.327	0
	9	3.8	2	2	1	4.956	0
	15	3.8	2	3	1	10.352	0
modality	mode	duration (s)	pause period (s)	# of repeats	n	mean (s)	std (s)
vibration only	7	1.8	2	2	1	3.586	0
	11	2.8	1	3	3	10.866	2.786
	14	2.8	2	3	1	8.443	0
	6	3.8	1	2	1	4.53	0
	12	3.8	1	3	1	6.964	0
	15	3.8	2	3	1	8.912	0
	3	3.8	n/a	1	1	3.054	0

However, a couple of observations regarding watch acknowledgement response times are noted:

(a) On the average across participants and received alarms, participants acknowledged alarms more when they received vibration + sound alarms and acknowledged them lesser when the modality is vibration only.
(b) On the average, participants tend to acknowledge alarms faster when they received shorter duration, vibration and sound alarms.
(c) On the average, participants acknowledged alarms slower when they received longer duration alarms with repeats.

As observed in Table 5, participants (n=7), on average, turned their heads the fastest (1.188 seconds) after vibration alarms of mode 6 (duration of 3.8 seconds duration, repeated every second, repeated two times) and turned their heads the slowest (n=3, 6.289 seconds) after vibration alarms of mode 14 (duration of 2.8 seconds duration, repeated every two seconds, repeated three times) across all configurations. However, when we look at modality 1 (sound and vibration) only, the fastest reaction time is with mode 13 with 1.52 seconds on the average for 1.8 sec duration with three repeats, whereas the slowest reaction time is with mode 3 with 3.43 seconds on the average for a long 3.8 sec duration. When we look at modality 2 (vibration only), it is the same finding for the entire table (mode 6 fastest, mode 14 slowest). The table further indicates that:

(a) regardless of duration and repetitions, it is the alarm modality people pay more attention to with vibration and sound combination people being more responsive to,
(b) longer durations with longer pause periods in alarms receive faster response from participants, and
(c) shorter durations with short/no pause periods with less repetitions receive faster response for **head turns** towards the direction of the traffic.

When we look at the vehicle detection times within the field of view of participants when they turn their heads towards the traffic direction, we observe the findings in Table 6. Participants (n=10) see hazardous vehicles the fastest (0.102 second) after sound and vibration alarms (duration of 1.8 seconds, repeated every 2 seconds, repeated three times) and slowest (n=7, 5.787 seconds) after vibration alarms (duration of 1.8 seconds duration, repeated every second, repeated three times). The table further indicates that:

(a) when vibration and sound are combined as the modality, people are quicker in detecting hazardous vehicles in shorter durations (1.8 and 2.8 seconds) as compared to 3.8 sec duration alarms,
(b) when vibration only alarms are sent, people show erratic response across durations, but tend to be quicker in detecting vehicles at shortest duration alarms with longer pause periods (1.8 seconds) or longest duration with repeats.

Table 5: Statistics for head turn reaction times in response to each alarm configuration

Alarm Characteristics					Head turn Reaction Time (s)		
modality	mode	duration (s)	Pause period (s)	# of repeats	n	mean (s)	std (s)
sound + vibration	7	1.8	2	2	10	1.937	1.885
	13	1.8	2	3	10	1.52	1.475
	4	1.8	1	2	10	3.194	2.314
	10	1.8	1	3	13	2.793	3.361
	1	1.8	n/a	1	5	2.42	1.823
	8	2.8	2	2	9	2.709	1.908
	14	2.8	2	3	9	3.266	2.308
	5	2.8	1	2	7	2.801	2.05
	11	2.8	1	3	2	3.229	0.451
	2	2.8	n/a	1	8	2.812	2.409
	9	3.8	2	2	7	2.839	2.231
	15	3.8	2	3	8	1.881	1.334
	6	3.8	1	2	13	1.659	1.392
	12	3.8	1	3	2	1.772	1.375
	3	3.8	n/a	1	7	3.43	4.228
modality	mode	duration (s)	Pause period (s)	repeats	n	mean (s)	std (s)
vibration only	7	1.8	2	2	10	2.374	3.361
	13	1.8	2	3	8	2.414	2.77
	4	1.8	1	2	8	3.871	3.765
	10	1.8	1	3	7	2.427	1.537
	1	1.8	n/a	1	3	5.233	2.998
	8	2.8	2	2	7	1.433	0.92
	14	**2.8**	**2**	**3**	**3**	**6.289**	**4.037**
	5	2.8	1	2	13	2.031	0.906
	11	2.8	1	3	9	2.759	1.618
	2	2.8	n/a	1	13	2.91	2.155
	9	3.8	2	2	6	2.404	3.152
	15	3.8	2	3	5	3.759	2.543
	6	**3.8**	**1**	**2**	**7**	**1.188**	**0.854**

	12	3.8	1	3	6	1.851	1.348
	3	3.8	n/a	1	15	1.969	1.449

Table 6: Statistics for vehicle detection times in response to each alarm configuration

Alarm Characteristics					Vehicle Detection Reaction Time		
modality	mode	duration (s)	pause period (s)	# of repeats	n	mean (s)	std (s)
sound + vibration	7	1.8	2	2	10	5.268	11.068
	13	**1.8**	**2**	**3**	**10**	**0.102**	**0.089**
	4	1.8	1	2	10	3.222	6.34
	10	1.8	1	3	13	1.034	1.344
	1	1.8	n/a	1	5	1.506	1.624
	8	2.8	2	2	9	1.895	2.428
	14	2.8	2	3	9	1.148	1.438
	5	2.8	1	2	7	1.664	2.196
	11	2.8	1	3	2	2.154	2.051
	2	2.8	n/a	1	8	1.741	2.176
	9	3.8	2	2	7	2.124	3.611
	15	3.8	2	3	8	2.019	1.652
	6	3.8	1	2	13	3.679	6.146
	12	3.8	1	3	2	5.021	4.945
	3	3.8	n/a	1	7	1.684	1.244
modality	mode	duration (s)	Pause period (s)	repeats	n	mean (s)	std (s)
vibration only	7	1.8	2	2	10	1.261	1.768
	13	1.8	2	3	8	1.23	1.507
	4	1.8	1	2	8	2.447	2.59
	10	**1.8**	**1**	**3**	**7**	**5.787**	**10.849**
	1	1.8	n/a	1	3	0.547	0.697
	8	2.8	2	2	7	1.777	2.194
	14	2.8	2	3	3	3.912	3.373
	5	2.8	1	2	13	2.012	2.029
	11	2.8	1	3	9	2.838	3.652
	2	2.8	n/a	1	13	0.973	1.684

	9	3.8	2	2	6	2.134	2.245
	15	3.8	2	3	5	0.289	0.252
	6	3.8	1	2	7	1.52	1.817
	12	3.8	1	3	6	0.773	0.95
	3	3.8	n/a	1	15	1.752	1.927

Regarding body movements, certain participants (n=6) on average, moved their bodies 1.048 seconds after receiving vibration alarms of mode 9 (duration of 3.8 seconds, repeated every 2 seconds, repeated twice) (Table 7). On the other extreme, participants (n=10) moved their bodies 4.214 seconds after receiving sound and vibration alarms of mode 13 (duration of 1.8 seconds, repeated every 2 seconds, repeated three times). The table also indicates that:

(a) when vibration and sound are combined as the modality, people are in general quicker in moving away from the traffic when they receive the alarms in longer durations, or alarms with shorter pause periods.
(b) when vibration only alarms are sent, people tend to be quicker with shortest duration alarms or longest duration alarms with long pause periods.

Table 7: Statistics for body reaction times in response to alarm characteristics

Alarm Characteristics					Body Movement Reaction Time		
modality	mode	duration (s)	Pause Period(s)	repeats	n	mean (s)	std (s)
sound + vibration	4	1.8	1	2	10	2.878	1.909
	10	1.8	1	3	13	3.254	2.579
	7	1.8	2	2	10	2.822	3.754
	13	**1.8**	**2**	**3**	**10**	**4.214**	**3.682**
	1	1.8	n/a	1	5	1.727	1.386
	5	2.8	1	2	7	2.152	2.238
	11	2.8	1	3	2	2.293	1.618
	8	2.8	2	2	9	3.659	3.551
	14	2.8	2	3	9	2.274	2.136
	2	2.8	n/a	1	8	4.031	4.258
	6	3.8	1	2	13	2.085	2.227
	12	3.8	1	3	2	2.676	0.348
	9	3.8	2	2	7	3.517	2.227
	15	3.8	2	3	8	2.515	2.325

	3	3.8	n/a	1	7	2.02	1.056

Table 7(continued): Statistics for body reaction times in response to alarm characteristics

Alarm Characteristics					Body Movement Reaction Time		
modality	mode	duration (s)	Pause (s)	repeats	n	mean (s)	std (s)
vibration only	4	1.8	1	2	8	2.579	1.758
	10	1.8	1	3	7	1.884	0.967
	7	1.8	2	2	10	2.036	1.283
	13	1.8	2	3	8	2.537	2.798
	1	1.8	n/a	1	3	2.047	0.552
	5	2.8	1	2	13	3.187	3.192
	11	2.8	1	3	9	2.744	2.653
	8	2.8	2	2	7	2.122	0.726
	14	2.8	2	3	3	3.304	2.059
	2	2.8	n/a	1	13	2.373	1.717
	6	3.8	1	2	7	2.232	1.085
	12	3.8	1	3	6	2.746	1.93
	9	**3.8**	**2**	**2**	**6**	**1.048**	**1.111**
	15	3.8	2	3	5	1.35	0.969
	3	3.8	n/a	1	15	3.587	3.565

When only looking at alarm modality, we observed that participants doubled their response in acknowledging sound + vibration alarms as compared to vibration only alarms. This means that combining multiple modalities in an alarm configuration could be more effective for worker reactions and should be further investigated. Mean acknowledgement times were also with a similar conclusion where participants were observed to be faster in responding to sound and vibration alarm configuration (i.e., with 6.129 seconds on the average) as compared to vibration only alarms (i.e., 7.565 seconds on the average) (Table 8).

Table 8: Statistics for watch acknowledgement times in response to alarm modality

Alarm Characteristics	Watch Acknowledgement Time			
modality	n_acknowledge	% acknowledge	mean	std
sound + vibration	17	14.2%	6.129	3.829
vibration only	9	7.5%	7.565	3.405

Table 8 was used for Monte Carlo Simulations and RL model simulations for predicting when and how fast workers would acknowledge the alarms on the smartwatch. In other words, only the modality of the alarm was accounted for when predicting worker reactions on the smartwatch in subsequent analyses.

Monte Carlo Reward Simulation

Using the reward function defined in the previous section and a Monte Carlo simulation was run for predicting the expected rewards from 1000 random alarm characteristic combinations based on the means and standard deviation worker reaction times in Tables 4-8. Table 9 presents the average rewards received given the configurations of alarms. Vibration alarms of mode 1 (duration of 1.8 seconds, never repeated) and 15 (duration of 3.8 seconds, repeated every 2 seconds, repeated three times) resulted in the highest rewards on average. Sound and vibration alarms of mode 1 (duration of 1.8 seconds, never repeated) and 3 (duration of 3.8 seconds duration, never repeated). This simulation findings overlap with the statistical results discussed in relation to Tables 4-7.

Table 9: Monte Carlo simulation based predicted-rewards for 1000 random alarm combinations

Mode	Modality	
	vibration	Sound + vibration
1	919.4	722.3
2	331.6	196.5
3	253.7	735.8
4	524.6	358.9
5	319.4	512.1
6	771.4	500.0
7	358.1	336.1
8	690.2	375.4
9	875.0	300.1
10	810.2	228.0
11	328.1	459.1
12	419.6	669.4
13	437.5	395.2
14	204.3	344.1
15	1000.0	699.1

Figure 25 illustrates how these alarm characteristics often result in the highest reward earned through a body movement reaction (R=1000) and less frequently resulting in no reaction at all (R=-250).

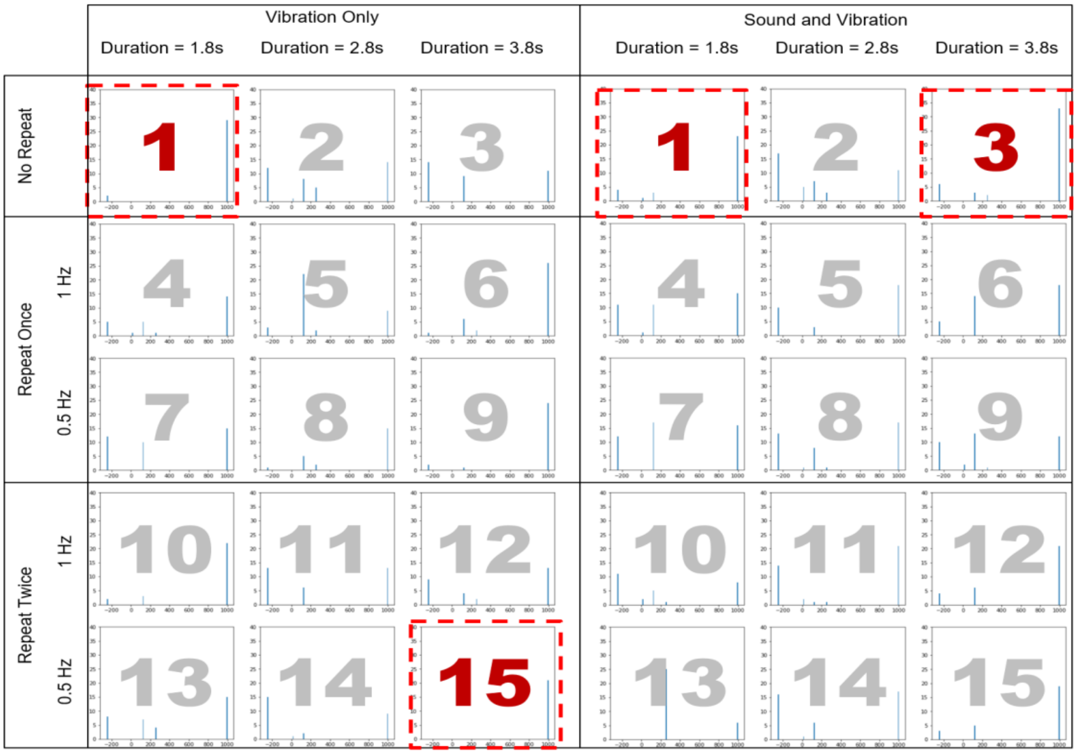


Figure 16: Histogram of rewards predicted by Monte Carlo Simulation of 1000 random alarm combinations

Reinforcement Learning Results

Using the backbone detailed in the research methodology section for RL model training, we trained an RL agent to configure and generate alarms that people are more responsive/attentive to. Figure 26 shows the alarm actions picked over the course of 700 episodes. Each dot represents a mode selected by the RL agent each time the SUMO simulation detected a hazardous vehicle. Red dots indicate vibration alarms while blue dots represent sound and vibration. Grey dots indicate when the RL agent did not raise the alarm when SUMO detected a hazardous vehicle. In general, the RL agent would take two to three alarm actions per episode before those episodes terminated when the synthetic worker completed placing all six cones, consistent with the VR user studies.

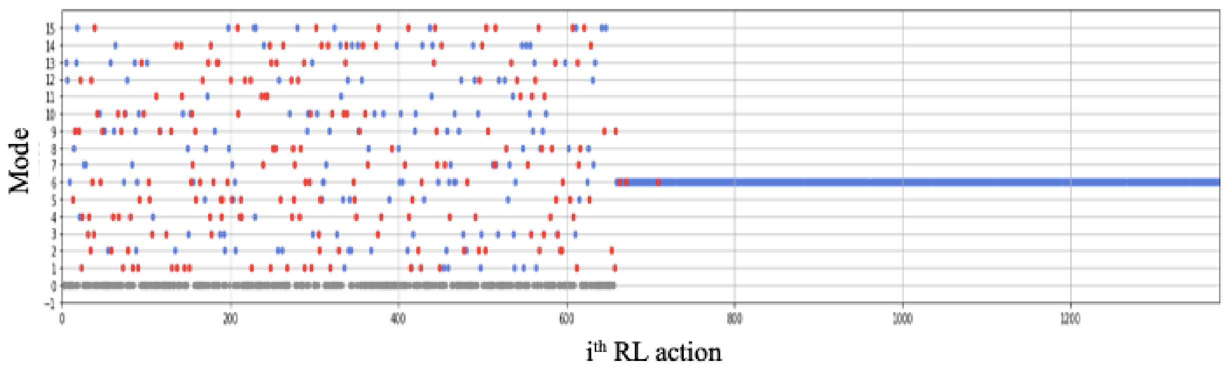


Figure 17: Actions performed by the RL agent over all training episodes. RL agent converged on sound + vibration alarm of mode 6 (duration of 3.8 seconds duration, repeated every second, repeated two times). Red dot: vibration only; Blue dot: Sound + vibration based alarms; Grey dot: RL agent skipped an alarm on a triggered event

As indicated in the Methodology section, the RL agent was defined to take random actions in the first 350 episodes before using the Proximal Policy Optimization in all subsequent episodes. Once the RL agent started using this optimization approach, it quickly settled on using sound and vibration alarms of mode 6 (duration of 3.8 seconds duration, repeated every second, repeated two times). Figure 27 below shows the moving average reward of every 100 actions taken by the alarm agent, with a grey line indicating the initial reward random actions taken in the first 350 episodes and a blue line indicating rewards resulting from the RL agent using the PPO approach. Rewards steadily increase then hover at around 500 at the end of the training episodes.

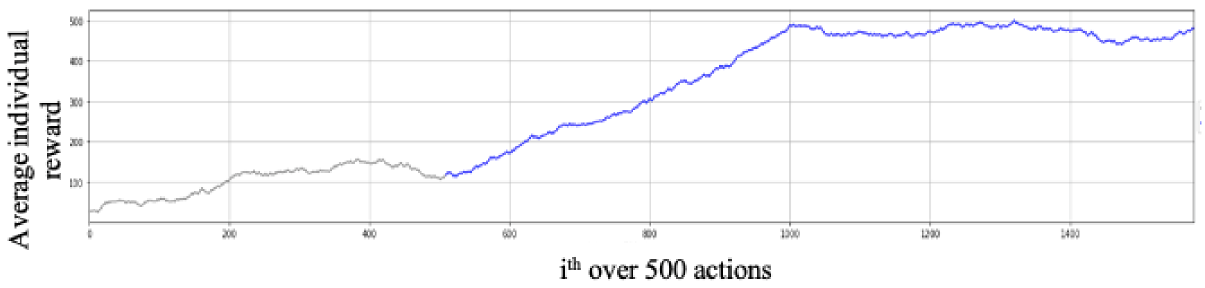


Figure 18: Moving average rewards over every 100 alarm actions by the RL agent. Blue line: RL actions; Grey line: Random actions

The first 350 episodes constituted a kind of Monte Carlo simulation for the RL agent to see a reward distribution for the full range of possible alarm characteristics before deciding on how to optimize. Table 10 below shows the average reward observed by the RL agent resulting from the alarm actions in the first 350 episodes, which can be compared to the Monte Carlo simulation results in Table 9. Sound with vibration alarms of mode 5 and vibration alarms of mode 6 (duration of 3.8 seconds duration, repeated every second, repeated two times) resulted in the highest average rewards after 350 training episodes. It is worth noting that *sound with vibration alarm mode 6*, the alarm action constantly selected by the RL agent towards the end of training showed a relatively small reward after the first 350 episodes, below the average reward observed from all alarm characteristic combinations (m=330.5)

Table 10: Average reward observed over the first 350 episodes where random alarm characteristics were raised by the RL agent

Mode	Modality	
	vibration	Sound + vibration
1	700.8	397.7
2	437.5	250.0
3	79.5	402.0
4	228.3	83.3
5	375.0	662.5
6	730.8	177.4
7	232.1	45.5
8	700.0	84.6
9	692.3	100.0
10	583.3	136.0
11	-145.8	-35.7
12	400.0	281.3
13	250.0	321.4
14	191.7	411.8
15	653.8	490.4

Discussion of Findings

Body and head reaction times averaged around 2.5 seconds. While the average watch acknowledgement time was far greater than 2 seconds, the most common watch acknowledgement time is also around 2 seconds. In general, these statistics from preliminary VR user studies are similar to literature figures citing reaction times to be around 2 seconds to alarms and other safety scenarios. While the current attention monitoring system does not distinguish between whether participants saw vehicles in their peripheral or center of vision, the general statistics indicate that vehicles are often already in a person's field of view during alarms. The fact that vehicle attention reaction times are much lower than body movement and head turn reaction times suggest that workers may have general awareness of vehicles well before they physically move to investigate the alarm's cause or avoid an accident. Generally, this brings into question whether safety systems should expect worker's situation awareness to change in the same expected amount of time as physical reactions. That said, the total number of participants in statistical analyses is 33 and much lower when looking at specific alarm characteristic combinations.

Monte Carlo simulation results show that vibration alarms of mode 1 (duration of 1.8 seconds duration, never repeated) and 15 (duration of 3.8 seconds duration, repeated every 2 seconds, repeated three times) as well as sound with vibration alarms of mode 1 (duration of 1.8 seconds duration, never repeated) and 3 (duration of 3.8 seconds duration, never repeated) resulted in higher relative rewards compared to other alarm characteristic combinations. These results are expected considering that these alarm modality and modes generally resulted in faster body movement reaction times during VR user studies and the reward function assigns the highest reward value for that form of worker reaction.

Surprisingly, the RL agent did not converge on any of these alarm characteristic combinations during the final training episodes. While its actions in the latter half of training did result in higher rewards on average, the RL agent also never seemed to actively explore the other alarm modes during its use of the PPO approach. Based on Table 10, the RL agent should have tried modes 1, 6, or 8 of the vibration alarm modality, since these modes resulted in the highest average rewards after 350 episodes. Furthermore, Figure 27 shows a "plateau" of rewards in the latter training episodes. Ideally, this plateau should have prompted the RL agent to explore other alarm characteristics to see they resulted in still higher rewards. Differences between Table 9 and 10 show that the RL agent could expect higher rewards from alarm characteristics different from Monte Carlo simulations. Table 10 shows that sound with vibration alarms modes 5 and 15 resulted in higher average rewards, while Table 9 shows modes 1 and 3 resulting in higher rewards for the same modality. It is possible that more randomized alarm exploration training episodes may be required for the RL agent to see the same expected reward distributions as Monte Carlo.

Outputs, Outcomes, Impacts

Outputs

Publications:

- Ergan, S., Zou, Z., Bernardes, S., Zuo, F., and Ozbay, K. (2022). "Developing an integrated platform to enable hardware-in-the-loop for synchronous VR, traffic simulation and sensor interactions." *Advanced Engineering Informatics,* 51, January 2022, 101476.
- Qin, J., Lu, D., and Ergan, S. (2022). "Towards increased situational awareness at unstructured work zones: Analysis of response data captured in VR based micro traffic simulations." *ICCCBE 2022 (abstract accepted).*
- Bernardes, S. D., Zou, Z., Zuo, F., Ergan, S., Khan, J. A. and Ozbay, K. (2021). "Development of a Virtual-Reality Based Immersive and Integrated Traffic Simulation Platform for Studying Traffic Work Zone Safety Problems". In TRB Annual Meeting, Transportation Research Board.

Posters:

- Lu, D., Bernardes, S., Zuo, F., Ergan, S., and Ozbay, K. (2022). "Urban ad-hoc construction zones: Human behavior evaluation towards safety notifications." Wagner, Urban Research Day, March 8, 2022, Manhattan, NY.

Exhibits:

- Lu, D., (2022). "Urban ad-hoc roadway construction zones: Worker behavior evaluation towards safety notifications." Tandon Research Excellence Exhibit, April 29, 2022, Brooklyn, NY.

Protoype:

- "Developing an integrated platform to enable hardware-in-the-loop for synchronous VR, traffic simulation and sensor interactions"- functional prototype developed for user studies and data collection on worker behavioral data during dangerous traffic simulations.
- 3 virtual reality models simulating three mobile/short term work zones and activities.

Dataset:

- Please check the data management plan for the submitted datasets for this project.

Code:

- Code developed in Phyton to get raw data captured in VR user studies (timestamped body movements) and convert them to identified safety metrics (i.e., # of times and time it takes for an alarm is acknowledged by a worker, # of times and time it takes for a worker to change gaze direction towards traffic flow, # of times and time it takes for a worker to detect hazardous vehicles after an alarm is triggered, # of times and time it takes for a worker to move away from the traffic flow).

- Code developed in Monte Carlo simulation to generate synthetic observation states for worker behaviors (which simulates a worker's response using historical worker responses towards generated alarms) to increase the training dataset for RL agent training.
- Code developed in Phyton to build the RL agent model and train the agent.

Outcomes

- The outcome of this research is an attempt to calibrate when, at what frequency, and how to (with what modalities) share warnings from simulated virtual environments as well as real vehicles and sensors with workers involved in active work zones for effective responses towards reduction of incidents.
- The developed VR enabled micro simulation platform serves as a platform to be leveraged for workforce training for increasing situational awareness at work zones.
- The reinforcement learning model rained from the user studies data can be further improved to deploy to physical work zones for alarm delivery to notify construction workers onsite for potentially dangerous situations.
- The research team initiated partnership conversations with Civil and Environmental (CEE) and Electrical and Computer Engineering (ECE) faculty members from North Carolina at Charlotte (NCC) for collaborating on data sharing and potentially submitting a joint NSF proposal.
- The research project's overview is presented online at: https://c2smart.engineering.nyu.edu/work-zone-safety-iii-calibration-of-safety-notifications-through-reinforcement-learning-and-eye-tracking/
- The research team created a database hosted at C2SMART's PostgreSQL server for storing the data collected in real-time by the sensors implemented for virtually delimiting the bounds of the work zone based on the physical environment of the lab. The database also serves as backend for the VR + Traffic simulation platform. The database is restricted access for now, only the research team has the credentials to insert and retrieve data from it.
- The research team hosted 1 high school student through the ARISE program at NYU.
- The research team hosted 2 undergraduate students (1 CUE and 1 CSE) to work with the research team on data analysis.

Impacts

- The research will produce a RL based models to implement in different real-word work zone scenarios with hazardous situation for workers and notify workers in maximum efficiency (i.e., optimum number of notifications without losing worker attention). Research outcomes will be shared with DOTs and FHWA and suggested as policy implementations for reduction of incidents at work zones.
- The research outcomes are aiming to reduce the number of incidents/accidents due to worker attention problems at work zones through reduction of alarm fatigue. In 2019, the Federal Highway Administration reported 762 crashes at roadway work zones, resulting in 842 fatalities, 135 of which were construction workers (BLS, 2019). Nearly a third of these fatalities involved

speeding vehicles (FHWA, 2022). A recent survey conducted by the Associated General Contractors (AGC) of America found that sixty percent of highway contractors reported motor vehicle accidents in their construction work zones during 2020 (AGC, 2021).

Conclusion

Overall, results indicate that the RL agent can at least perform better than random alarm actions after a small number of training episodes. The optimization approaches within the field of RL research (PPO, actor-critic, Deep Q-learning) will need to be explored further to see if the RL agent can at least choose alarm actions as expected by Monte Carlo simulations. Specifically, this future work constitutes investigating the Tensorforce code library and other machine learning code bases for ways to effectively explore and optimize alarm actions. The "plateau" of rewards can also be the result of how the reward function is specified. The implemented reward function assigns a constant value based on which worker reaction to the alarm is the fastest, regardless of how short the worker's reaction time is and regardless of whether the worker quickly reacted in a combination of forms (e.g., head turn and body movement). Different reward functions that can produce more variation in reward values could possibly provide a better reward gradient for the RL agent to search over. For example, the reward function can also increase the reward for body movements if the worker's body movement reaction time is one or more standard deviations smaller than the expected body move reaction time. Improvements to the attention monitoring system, such as distinguishing between when workers see vehicles in their central field of view or just their periphery, can also feed into ways the reward function can vary the final reward value for a better gradient. Differences observed between rewards resulting from randomized RL actions and predicted by Monte Carlo simulations suggest that RL agents will need a significantly more training episodes to potentially see the full "landscape" of rewards resulting from all alarm actions. These improvements to the RL agent reward function, optimization approach, and training configurations will help future work on using RL to optimize alarm characteristics during live VR user studies.

References

1. AGC (2021). "Associated General Contractors of America -Sixty Percent Of Firms Working On Highway Upgrades Experienced Cars Crashing Into Their Work Zone During The Past Year, New Data Finds." Available at: https://www.agc.org/news/2021/05/27/sixty-percent-firms-working-highway-upgrades-experienced-cars-crashing-their-work (accessed June 27, 2021).
2. Akanmu, A., Olayiwola, J., Ogunseiju, O., and McFeeters, D. (2020). "Cyber-physical postural training system for construction workers." *Automation in Construction*, 117 (9), 103272, doi: 10.1016/j.autcon.2020.103272.
3. AWARE (2020). "Asphalt Pro. Oldcastle's AWARE System Makes Every Second Count." Available at: https://theasphaltpro.com/articles/oldcastle-aware-system/. (Accessed July 29, 2020).

4. Bella, F. (2005). V"alidation of a driving simulator for work zone design." *Transportation Research Record*, 1937(1), 136-144.
5. Bureau of Labor Statistics (2019), "2019 National Work Zone Fatal Crashes & Fatalities," Available at: https://www.workzonesafety.org/crash-information/work-zone-fatal-crashes-fatalities/ (accessed November 27, 2021).
6. Burke, M. J., Salvador, R. O., Smith-Crowe, K., Chan-Serafin, S., Smith, A., and Sonesh, S. (2011). "The dread factor: how hazards and safety training influence learning and performance." *Journal of Applied Psychology*, 96(1), 46.
7. Domenichini, L., La Torre, F., Branzi, V., and Nocentini, A. (2017). "Speed behavior in work zone crossovers. A driving simulator study." *Accident Analysis & Prevention*, 98, 10-24.
8. DOT(2020). "Department of Transportation, Federal Highway Administration – FHWA. Work zone guides and documents (Listed by Title)" Available at: https://ops.fhwa.dot.gov/wz/outreach/wz_training/wz_guides_documents.htm. (Accessed July 24, 2020).
9. Du, J., Zou, Z., Shi, Y., and Zhao, D. (2018). "Zero latency: Real-time synchronization of BIM data in virtual reality for collaborative decision-making." *Automation in Construction*, 85, 51-64.
10. Ergan, S., Zou, Z., Bernardes, S., Zuo, F., and Ozbay, K. (2022). "Developing an integrated platform to enable hardware-in-the-loop for synchronous VR, traffic simulation and sensor interactions." *Advanced Engineering Informatics,* 51(1), 101476.
11. FWHA(2022). "Federal Highway Administration Work Zone Facts and Statistics," Available at: https://ops.fhwa.dot.gov/wz/resources/facts_stats.htm (accessed March 27, 2022).
12. Gheisari, M., and Esmaeili, B. (2019). "Applications and requirements of unmanned aerial systems (UASs) for construction safety." *Safety Science*, 118, 230-240. https://doi.org/10.1016/j.ssci.2019.05.015.
13. Intellicone (2020). "Trans Canada Traffic Inc. Intellicone." Available at: http://www.transcanadatraffic.ca/IntelliconeProductInformation.html#.WroA10xFzop. (Accessed July 24, 2020).
14. Jeelani, I., Asadi, K., Ramshankar, H., Han, K., and Albert, A. (2019). "Real-world mapping of gaze fixations using instance segmentation for road construction safety applications." *ArXiv*:1901.11078 [Cs, Stat]. http://arxiv.org/abs/1901.11078
15. Lee, D., and Kim, M. (2021). "Autonomous construction hoist system based on deep reinforcement learning in high-rise building construction." *Automation in Construction,* 128(8), 103737, doi: 10.1016/j.autcon.2021.103737.
16. Lee, W., Lin, K.-Y., Seto, E., and Migliaccio, G. C. (2017). "Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction." *Automation in Construction*, 83, 341–353. https://doi.org/10.1016/j.autcon.2017.06.012
17. Lewis, R. M. (1989). "Work-zone traffic control concepts and terminology." *Transportation Research Record*, 1230, 1-11.
18. Li, X., Yi, W., Chi, H. L., Wang, X., and Chan, A. P. (2018). "A critical review of virtual and augmented reality (VR/AR) applications in construction safety." *Automation in Construction*, 86, 150-162.
19. Lin, K. Y., Son, J. W., and Rojas, E. M. (2011). "A pilot study of a 3D game environment for construction safety education." *Journal of Information Technology in Construction* (ITcon), 16(5), 69-84.

20. Lucas, J., Thabet, W., and Worlikar, P. (2008). “A VR-based training program for conveyor belt safety.” *Journal of Information Technology in Construction (ITcon),* 13, 381-407, http://www.itcon.org/2008/25
21. Luo, X., Li, H., Huang, T., and Rose, T. (2016). “A field experiment of workers’ responses to proximity warnings of static safety hazards on construction sites.” *Safety Science,* 84, 216–224. https://doi.org/10.1016/j.ssci.2015.12.026
22. McAvoy, D. S., Schattler, K. L., and Datta, T. K. (2007). “Driving simulator validation for nighttime construction work zone devices.” *Transportation Research Record*, 2015(1), 55-63.
23. Mishra, S., Golias, M., and Thapa, D. (2021). “Work Zone Alert Systems”, Final Report No. RES2019-01, The University of Memphis. Available at:https://www.tn.gov/content/dam/tn/tdot/long-range-planning/research/final-reports/res2019-final-reports/RES2019-01_Final_Report_Approved.pdf (Accessed March 27, 2022).
24. Ng, A., Coates, A., Diel, M., Ganapathi, V., Schulte, J., Tse, B., Berger, E., and Liang, E. (2006). “Autonomous Inverted helicopter flight via reinforcement learning,” *Experimental Robotics IX*, vol. 21, M. H. Ang and O. Khatib, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 363–372. doi: 10.1007/11552246_35.
25. Park, J. and S. Sakhakarmi. (2019). “Embedded safety communication system for robust hazard perception of individuals in work zones.” Report for *CPWR-The Center for Construction Research and Training*, Available at: https://www.cpwr.com/wp-content/uploads/publications/SS2019-embedded-safety-work-zones.pdf (Accessed March 27, 2022).
26. Perlman, A., Sacks, R., and Barak, R. (2014). “Hazard recognition and risk perception in construction.” *Safety Science*, 64, 22-31. https://doi.org/10.1016/j.ssci.2013.11.019.
27. Sakhakarmi, S., and J. Park. “Investigation of tactile sensory system configuration for construction hazard perception.” *Sensors*, 19(11), 2527, 2019.
28. Santos, S. B., Givigi, S. N., and Nascimento, C. L. (2013). “Autonomous construction of structures in a dynamic environment using Reinforcement Learning,” *IEEE International Systems Conference (SysCon),* Orlando, FL, Apr. 2013, 452–459. doi: 10.1109/SysCon.2013.6549922.
29. Schrittwieser, J., Antonoglou, I., Hubert, T., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., Lillicrap, T., and Silver, D. (2020). “Mastering Atari, Go, chess and shogi by planning with a learned model,” *Nature*, 588(7839), 604–609, doi: 10.1038/s41586-020-03051-4.
30. SonoBlaster (2020). “SonoBlaster work zone intrusion alarm.” Available at: https://www.transpo.com/roadshighways/safety-products/wz-intrusion-alarm. (Accessed July 24, 2020.)
31. Tang, L., Rosales, R., Singh, A., and Agarwal, D. (2013). “Automatic ad format selection via contextual bandits.” in *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management - CIKM ’13*, San Francisco, California, USA, 1587–1594. doi: 10.1145/2505515.2514700.
32. Tapan D., Gates T., Savolainen P., Kay J., Parajuli S. and Nicita N. (2016). “A guide to short-term stationary, short-duration, and mobile work zone traffic control.” Report submitted to FHWA, Detroit, MI.
33. TRB (2005). “A guide for reducing work zone collisions.” Transportation Research Board National Academies of Sciences, Engineering, and Medicine – NASEM, Washington, DC. (2005).

34. Tuttas, S., Braun, A., Borrmann, A., and Atilla, U. (2017). "Acquisition and consecutive registration of photogrammetric point clouds for construction progress monitoring using a 4D BIM." *Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 85(1), 3-15. https://doi.org/10 .1007/s41064-016-0002-z.
35. USDOT (2009). "US Department of Transportation manual on uniform traffic control devices; for streets and highways." US Department of Transportation, Federal Highway Administration.
36. Vinyals et al. (2019). "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, doi: 10.1038/s41586-019-1724-z.
37. Wong, J. M., Arico, M. C., and Ravani, B. (2011). "Factors influencing injury severity to highway workers in work zone intrusion accidents." *Traffic Injury Prevention*, 12(1), 31-38.
38. Wu, C., Kreidieh, A., Parvate, K., Vinitsky, E., and Bayen, A. (2022). "Flow: A Modular Learning Framework for Mixed Autonomy Traffic." *IEEE Trans. Robot*., 38(2), 1270–1286, doi: 10.1109/TRO.2021.3087314.
39. Yang, H., and Lu, Q. (2016). "Dynamic contextual multi arm bandits in display advertisement," *IEEE 16th International Conference on Data Mining (ICDM)*, Barcelona, Spain, 1305–1310. doi: 10.1109/ICDM.2016.0177.
40. Zou, Z., Yu, X., and Ergan, S. (2020). "Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network," *Building and Environment*, 168(1), 106535, ,doi: 10.1016/j.buildenv.2019.106535.
41. Zuluaga, C. M., Namian M. and Albert A. (2016). "Impact of training methods on hazard recognition and risk perception in construction." *Construction Research Congress*, 2861–2871. Reston, VA: ASCE. https://doi.org/10.1061/9780784479827.285