Automation in Construction 169 (2025) 105876

Contents lists available at ScienceDirect

Automation in Construction

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

Impact of selective environmental sound attenuation on operator performance, stress, attention, and task engagement in teleoperated demolition

Patrick Borges Rodrigues [a], Burcin Becerik-Gerber [a], Lucio Soibelman [a,*], Gale M. Lucas [b], Shawn C. Roll [c]

[a] Sonny Astani Department of Civil and Environmental Engineering, Viterbi School of Engineering, University of Southern California (USC), Los Angeles, CA 90089, USA
[b] USC Institute for Creative Technologies, USC, Los Angeles, CA 90094, USA
[c] Chan Division of Occupational Science and Occupational Therapy, USC, Los Angeles, CA 90089, USA

ARTICLE INFO

Keywords:
Demolition
Electrodermal activity
Environmental sound
Heart rate variability
Performance
Robot
Safety
Stress
Teleoperation

ABSTRACT

The noise produced in demolition sites can mask safety-critical sounds that inform operators about task conditions and hazards. These problems are exacerbated in teleoperated demolition, where the separation between operator and site compromises operators' situation awareness and cognitive loads. This paper assessed the effects of environmental sounds with and without attenuation on the operators' performance and response (e.g., stress, attention, task engagement) during teleoperated demolition. Eighty participants completed three virtual demolition tasks under different environmental sound conditions, i.e., no sound (NS), unfiltered sound (US), and filtered sound (FS) with 20-dB attenuation of background noise and robot's sounds to allow focus on safety and task conditions. The results show that US induced more stress than NS and FS. Also, FS resulted in fewer collisions, faster reaction times, and greater attention and task engagement than US. These results can support the design of sound feedback interfaces for teleoperation in construction.

1. Introduction

Sound is crucial for communications, operations, and safety in construction, providing workers with information about site conditions, the operational status of tools and equipment, the proximity of obstacles, and the presence of hazards [1,2]. However, the excessive noise conditions typical of most construction sites can mask safety-critical sounds (e.g., reverse alarms, workers' spoken warnings), diminishing their effectiveness in alerting workers to hazards on site and issues with the machines, which has led to numerous accidents [3,4]. Beyond compromising auditory perception, construction noise can adversely affect construction workers' psychological and physiological responses, such as inducing changes in heart rates (HR) and stress [5] and creating distractions [6], which have further implications on performance, safety, and well-being. These noise-related challenges are especially critical in teleoperation applications. The separation between operators and the site can compromise operators' perception of sensory information, such as sound, reducing situation awareness (SA) and increasing cognitive load [7,8]. As a result, operators may be less aware of the presence, location, and relative distance of hazards and obstacles in teleoperation, increasing the risk of accidents involving teleoperated machines.

While noise is a challenge in traditional on-site operations and teleoperation, the nature of these challenges differs in these applications. On-site operators have access to direct sensory feedback (e.g., visual, sound, haptics), which helps them to assimilate and respond to site conditions more naturally [9]. Alternatively, operators in teleoperation applications are physically separated from the site and must rely on mediated sensory feedback from teleoperation interfaces. This separation can distort or diminish critical sound cues. As a result, even moderate noise levels can have a more pronounced effect on teleoperators, as they lack the rich sensory context available to on-site operators.

Teleoperated machines are becoming increasingly common on construction sites, particularly in hazardous environments where operator safety is a significant concern [10]. Despite their growing use, the sound feedback systems of these teleoperated machines are often simplistic,

* Corresponding author at: Dept. of Civil and Environmental Engineering, Viterbi School of Engineering, Univ. of Southern California, KAP 224A, 3620 South Vermont Ave., Los Angeles, CA 90089-2531, USA.
E-mail address: soibelman@usc.edu (L. Soibelman).

https://doi.org/10.1016/j.autcon.2024.105876
Received 30 July 2024; Received in revised form 25 October 2024; Accepted 11 November 2024
Available online 23 November 2024
0926-5805/

typically comprising microphones and basic audio interfaces (e.g., [9,11]). These limited sound systems may fail to capture the complete auditory environment of the construction site, making it difficult for operators to perceive safety-critical sounds and further exacerbating the challenges posed by excessive noise.

Teleoperated demolition is particularly susceptible to these issues due to the high noise levels generated during demolition tasks. Demolition sites experience some of the highest noise levels in construction, reaching up to 120 A-weighted decibels (dBA) [12], far exceeding the National Institute for Occupational Safety and Health (NIOSH) recommended exposure limit of 85 dBA averaged over an eight-hour workday [13]. While teleoperation interfaces allow demolition operators to adjust the sound volumes they experience, the high noise levels on-site can still mask safety-critical sounds that alert operators to the risks of collapses and approaching equipment and workers. One potential solution is to use sound filters to attenuate excessive sounds and background noise, thereby improving the perception of safety-critical sounds generated on site [14].

In this paper, we investigate the effects of selective environmental sound attenuation on operators' task performance and safety in teleoperated demolition. We attenuated background noise and excessive noises from the demolition robot's engine, motion, and end-effector (breaker) while leaving safety-critical sounds, such as voices of approaching workers, equipment, and collapses, unattenuated. Considering that construction noise can also affect operators' psychological and physiological responses (e.g., heart rate, electrodermal activity, workload, SA, stress, attention, task engagement) [1,5], we also examined the impact of selective sound attenuation on operators' workload, SA, stress, attention, and task engagement. In an experimental study using a virtual environment, we compared operators' demolition performance and responses under three sound conditions: No Sound (NS) (baseline), Filtered Sound (FS) with 20 dB attenuation of background noise and robot sounds, and Unfiltered Sound (US). The study was designed to answer the following research questions:

1. How do the operators' task and safety performance in teleoperated demolition change with the inclusion of a selective sound attenuation filter to the interface compared to the interface providing unfiltered sound?
2. How do the operators' response: stress, attention, and task engagement in teleoperated demolition change with the inclusion of a selective sound attenuation filter to the interface compared to the interface providing unfiltered sound?

The remainder of this paper is organized as follows. Section 2 reviews related studies, and Section 3 reviews current gaps and opportunities for sound feedback in teleoperation interfaces. Section 4 presents a description of the research methodology. Section 5 presents the results of the experiment. Section 6 presents a discussion of the main findings, the limitations of the study, and future research opportunities. Finally, section 7 presents the conclusions of this study.

2. Literature review

2.1. Effects of sound feedback on the operators' performance and psychological and physiological responses in teleoperation

Sound feedback is crucial in teleoperation applications because it can improve operators' task performance, safety performance, and psychological and physiological responses. These effects are commonly associated with the increased SA and reduced mental workload that sound feedback provides in some cases [15]. Regarding task performance, sound feedback can increase position and orientation accuracy in manipulation tasks [16] and reduce reaction times compared to no sound feedback [17]. Additionally, sound feedback informs operators about the robot's operational status, movements, and potential malfunctions [18,19]. Regarding safety, sound feedback can provide proximity alerts that reduce the risk of collisions [20,21] and inform operators about the presence and levels of toxic and radioactive substances on site [15], even when the hazards are not in the cameras' fields of view.

Moreover, sound feedback positively affects operators' mental states and physiology in teleoperation applications. Many studies have shown positive psychological effects of sound feedback, including increased feelings of presence [22] and agency [19], enhanced SA [21], and reduced cognitive load [15] and distractions [23]. By improving operators' attention, alertness, and reaction times, sound feedback can serve as a redundant sensory stimulus to alert operators when they are fatigued and less attentive to visual information [24]. Many of these positive effects result from the reduced cognitive load achieved by distributing information across multiple sensory channels (e.g., vision, sound, haptics) [25].

However, improperly designed sound feedback interfaces and exposure to high noise levels can negatively affect operators' performance and psychological and physiological responses in teleoperation. Noisy environments affect completion times and distance traveled by the robots' end-effectors, effects that are worsened in more complex tasks [26]. Additionally, noise can compromise communications [1] and decision-making [27] and create distractions that result in safety performance degradation [6]. Noise levels and types also influence operators' cognitive loads and overall comfort. Various studies reported that operators felt uncomfortable [17,28] or experienced increased heart rates and stress [5,15] for higher noise levels and exposure durations, which indicates the need to control the sound levels operators experience during teleoperation.

3. Knowledge gaps

Although the studies in Section 2.1 show many benefits of sound feedback in teleoperation interfaces on operators' performance and psychological and physiological responses, most of these studies focused on the effects of synthetic sounds produced using data sonification instead of environmental sounds (e.g., [15,20]). While data sonification can be a powerful tool in informing operators about the presence, levels, and distances of hazards on site, it still faces many challenges compared to environmental sounds. Understanding synthetic sounds from data sonification can be more cognitively taxing for operators, as it requires inferential processes and mapping the synthetic sounds to their intended meanings, whereas understanding environmental sounds occurs more naturally [29]. Additionally, creating the data representations for sonification is a challenging task that requires domain knowledge since improper auditory representations can lead to increased workload [15].

Alternatively, while some studies considered environmental sounds in construction applications, most did not focus on robot teleoperation. Instead, they examined how construction sounds and noise affect operators' performance, safety, and well-being during traditional construction tasks simulated in virtual reality (VR) environments (e.g., [28,30]) or shown in images (e.g., [6]). These studies show that, in traditional construction applications, environmental sounds and construction noise affect operators' task and safety performance [31] and psychological and physiological responses [5]. However, although results from traditional construction applications can help inform the design of teleoperation interfaces, key differences exist between traditional and teleoperated construction tasks, particularly in how operators control the robots and assimilate sensory information from the site [8]. Thus, further characterization of the effects of environmental sound on operators' task and safety performance and psychological and physiological responses is needed in construction teleoperation applications.

In commercial teleoperated machines used in construction applications, environmental sound is captured by microphones installed in the cabins of the machines and transmitted to the operators via the teleoperation interface (e.g., [9,11]). However, these setups lack spatial audio capabilities, which may hinder operators' ability to perceive the direction and distance of safety-critical sounds (e.g., approaching machines, collapses), potentially impacting operators' situation awareness. Research efforts have explored the feasibility and usefulness of spatial audio and advanced audio processing techniques to support sound source localization (SSL) and increase operators' situation awareness in teleoperation [32]. Despite these potential benefits, spatial audio technology has not been used in commercial construction teleoperation systems.

As many studies show, construction sites are characterized by many sound sources (i.e., polyphonic settings) and high noise levels, which complicate understanding audio signals even when computational models are used [33]. The inability to perceive environmental sounds at appropriate levels due to hearing loss and noisy site conditions is a cause of concern among workers due to its potential impacts on job safety, stress, fatigue, communications, and the ability to hear warnings and equipment [34]. Similarly, noisy conditions on construction sites compromise communications [1] and can lead to accidents [3,4]. Despite that, insufficient research exists on augmenting construction workers' perception of safety-critical sounds generated on site [14].

Detecting anomalous sound events and enhancing speech communication in noisy environments have become areas of increased research interest in many fields. Machine learning (ML) models based on audio data have been used to detect screams and gunshots in surveillance applications, road crashes in traffic monitoring, and fault equipment in industrial settings [35]. In construction applications, ML models based on audio data from construction sites have been used for construction task monitoring [33], equipment activity recognition [36], infrastructure damage detection [37], hazard detection [2], and improved speech communication [38], among others. In all these cases, ML models effectively detected and alerted workers about potential hazards or problems on site. Opportunities exist to use advanced sound filtering that reduces the influence of unwanted background noise in traditional construction applications, which can support workers in detecting safety-critical sounds on site [14]. However, understanding the effects of these ML models and filters to enhance workers' perceptions of hazardous sounds from the site remains largely unexplored in construction teleoperation applications.

4. Methods

4.1. Use case

This study focused on a simulated teleoperated demolition task in a virtual environment using a Brokk 110 demolition robot (Figs. 1 (a) and (b)). During training and subsequent demolition task scenarios, the participants were asked to drive and position the body of the robot, its outriggers, and its arms to demolish walls and slabs using Brokk's hydraulic breaker. Although the proposed intervention is intended for real demolition robots, the use of a virtual environment in this study was justified based on the significantly lower costs, complexity, and risks and the relatively greater ability to control the experimental conditions (i.e., sound levels) that can be achieved in a virtual environment compared to a real construction site.

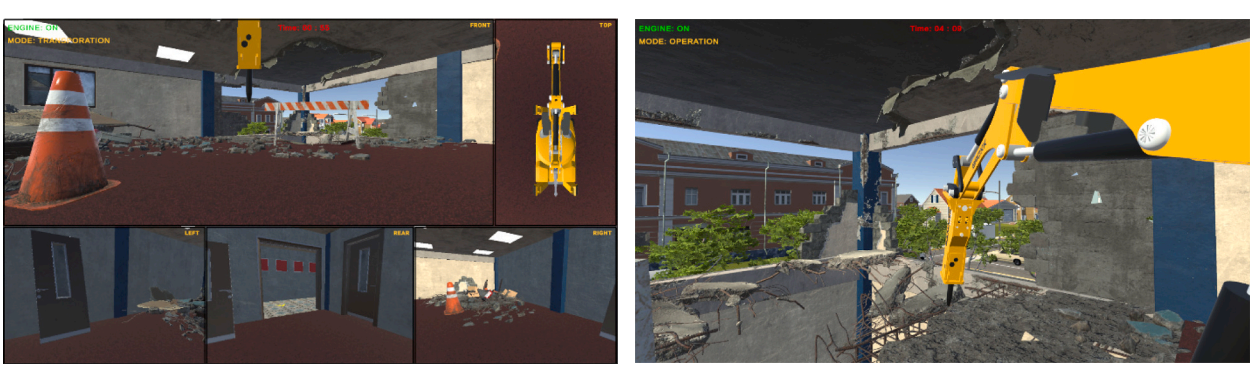
Demolition is one of the most dangerous construction activities in which workers are exposed to various safety and health hazards, such as high noise levels, dust, and collapses. Even when operators remotely control these robots using a control box while standing a few meters away from the robot, operations can be dangerous. In many cases, the safety and health risks associated with hazardous materials, explosions, noise, falling debris, or structural failures have resulted in many injuries and deaths among demolition workers [39,40]. Regarding exposure to construction noise, previous investigations have shown that demolition workers can experience high noise levels in construction sites, which can be over 120 dBA [12].

4.2. Workstation

The workstation configuration (Fig. 2) included a laptop powered by an Intel® Core™ i7-11800H, with 32 GB of RAM and 16 GB GPU (NVIDIA GeForce RTX 3080). The setup also included a 27-in. computer monitor, two ambidextrous joysticks (Thrustmaster T.16000 M Space Sim Duo Stick), and an immersive cinematic 3D audio headset (Audeze Mobius Spatial Audio Gaming Headset). To ensure comfort and functionality, the laptop and monitor were mounted on an adjustable desk, and the joysticks were attached to the arms of an ergonomic office chair.

4.3. Virtual environment design

The virtual environment was developed using the Unity game engine and aimed to create a highly realistic and interactive simulation for teleoperated demolition. We used realistic asset packages, such as the Destroyed Building Kit (Loknar Studio), to model the architectural and structural components of the building. These modular elements, including intact and destroyed versions, were assembled to represent different building regions where demolition tasks occurred (See Figs. 1 (a) and (b)). This modular design allowed us to control levels of destruction for each element by replacing their meshes and materials, enhancing the model's realism. To simulate the destruction of the building components during demolition, we used the RayFire for Unity plugin (RayFire Studios), a tool commonly employed in visual effects and game development for dynamic simulations of objects' destruction. RayFire enabled the realistic fragmentation, shattering, and collapse of building components, closely mimicking how materials would behave during demolition. In addition to RayFire, the physical interactions between the robot and the building elements—such as gravitational forces



(a) **(b)**

Fig. 1. (a) Transportation mode and (b) operation mode visual interfaces of the teleoperated robot.

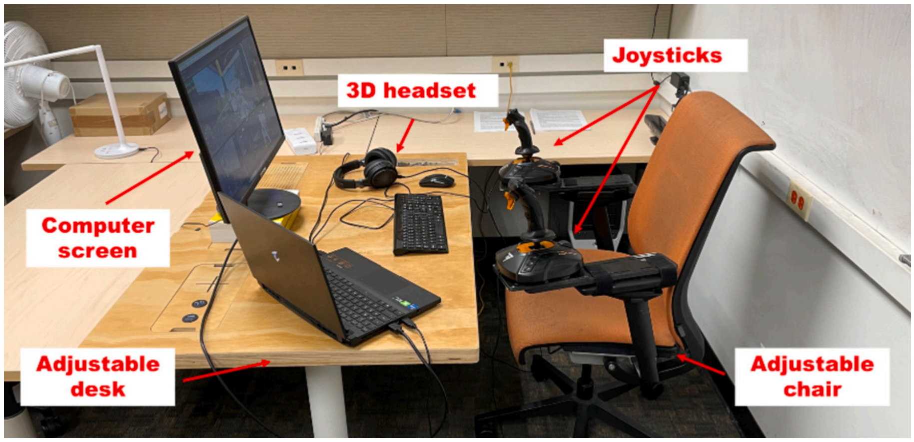


Fig. 2. Teleoperation workstation used in the experiment.

acting on the robot and built elements, collisions, end-effector forces applied during demolition, and interactions between concrete and rebar—were simulated using Unity's built-in physics engine and custom scripts.

The control system for the virtual robot replicated the control mechanisms of the real demolition robot, with joystick inputs mapped directly onto the virtual robot's arm joints, slew system, outriggers, and breaker. The system was designed to handle two operational modes: transportation and operation. As with the real robot, the transportation mode was used to control the robot tracks and outriggers. Alternatively, the operation mode was used to control the robot's arm, slew system, and end-effectors. The system dynamically adjusted camera perspectives based on the mode, providing operators with different views for transportation and operation (See Figs. 1(a) and (b)). To enhance the teleoperation interface, we carefully replicated the sensor placement on the robot body, including cameras and microphones, to mirror real-world constraints in teleoperation setups used in construction (e.g., [9,11]).

To further enhance the immersiveness and realism of the virtual environment, we used the Steam Audio Spatializer plugin combined with a spatial audio headset to simulate spatial audio. It allowed operators to accurately locate sound sources based on their position in the virtual environment. It also dynamically adjusted sound levels as the relative distances between sound sources (e.g., robot, equipment, workers) and the robot's microphone changed. Audio clips for the scene elements (e.g., robot sounds, worker voices, equipment) were extracted from actual construction and demolition operations recordings to increase sound realism. The system also maintained the relative loudness of these sounds, with louder noises (e.g., robot engine, breaker operation) masking quieter ones (e.g., footsteps, worker voices), replicating real-world conditions. For more details on the sound design for the experimental conditions, see Section 4.4.3. Experimental Conditions. Further information on the development of the virtual environment and its sound conditions is available in [41].

4.4. Study procedures

The study included three stages: (1) baseline measurements and training in the virtual environment, (2) completing a demolition task in a virtual environment with no sound as a baseline condition, and (3) completing a new demolition task in two experimental conditions using unfiltered and filtered sound. The order of the two experimental conditions was randomized to control for order effect on the outcome measures. Each demolition scenario lasted at most 10 min, and, for most participants, the experimental procedure lasted about 1 h and 15 min. USC's Institutional Review Board (IRB) approved the study protocol and its associated methods and procedures for surveys, questionnaires, psychometric tests, and psychological and physiological tests (Study ID UP-22-00914).

4.4.1. Baseline measures and training

After signing a consent form, the participants completed a demographic survey that included questions about gender, age, language, race/ethnicity, level of education, previous experiences at construction sites, in operating construction machines/robots, with video games, and virtual reality. Participants completed a short version of the Spielberger State-Trait Anxiety Inventory (STAI) consisting of 6 items [42] (See Section 4.6.2. Participants' Response). Then, participants were trained on the control mechanisms of the virtual demolition robot, including basic operations of the robot's tracks, arms, outriggers, and end-effectors, as well as on identifying and reacting to safety hazards. The training session was conducted on-screen using the simulation to provide instructions on operating the robot and behaving in unsafe situations without the researchers' intervention. Following the training, baseline physiological response data (i.e., electrodermal activity (EDA) and heart rate (HR)) was collected while the participants were asked to relax and remain still for five minutes. Then, participants completed the 6-item STAI assessment again.

4.4.2. Baseline condition

To complete all demolition tasks, the participants used joysticks to drive the robot and position its arms, outriggers, and end-effectors to perform the demolition task. The first demolition task for all participants was performed with no sound (NS). This condition was selected as the baseline condition. In the NS condition, participants demolished one wall and one suspended slab. Physiological response data (i.e., EDA and HR) was collected throughout the task, and, at the end of it, participants completed the 6-item STAI and the NASA Task Load Index (TLX) [43] assessments (See Section 4.6.2. Participants' Response).

4.4.3. Experimental conditions

Two experimental conditions were tested: unfiltered sound (US) and filtered sound (FS). In the US condition, the participants could hear various sounds from the simulated construction site, including the engine and movements of the robot, the breaker, traffic, construction equipment, workers, collapses, and background noise. These sounds were extracted from videos and audio clips recorded at real demolition sites. Each sound was associated with specific robot motions (e.g., arm, tracks, outriggers movements), operational conditions (e.g., engine on, breaker activated), or elements of the site (e.g., workers, equipment, traffic). In the FS condition, the sound sources associated with the engine and movements of the robot, the breaker, traffic, and background noise were attenuated, whereas the specific sounds from construction equipment, workers, and collapses remained unattenuated. This design of the FS condition differed from simple sound attenuation because not all sounds were attenuated. Instead, some sounds were attenuated while the sounds of interest stood out (i.e., selective environmental sound attenuation).

The selected sound conditions were designed to replicate real-world auditory challenges in teleoperated demolition and to evaluate potential solutions. Although teleoperation allows operators to control the sound volumes they experience, demolition sounds' complex and polyphonic nature makes it difficult to identify critical auditory cues. Moreover, recent studies have highlighted the need for mechanisms to enhance the perception of critical sounds in noisy construction settings [14]. Therefore, the US condition replicated traditional teleoperated demolition sound interfaces, while the FS condition served as an alternative sound interface that enhances sound perception by selectively attenuating non-critical and excessive sounds.

To control the sound levels the participants experienced during the simulations, audio mixers in Unity and the 3D headset were used. In Unity, the threshold volume of the audio mixers for both the US and FS conditions was set to 80 dB. Additionally, the volume of the 3D headset was set so that the average dBA levels experienced by the participants was around 60 dBA throughout most of the experiment but no more than 80 dBA for any combination of sounds. The average 60 dBA level was selected because it is the usual sound level one experiences during normal conversations [44]. Alternatively, the chosen limit of 80 dBA level is lower than the recommended exposure limit of 85 dBA (averaged over an eight-hour workday) set by NIOSH [13]. The dBA levels were measured using a BAFX3370 decibel meter [45]. The only difference between the FS and US conditions was that, in the FS condition, a 20 dB attenuation filter was included in the audio mixer associated with the attenuated sound sources (i.e., the engine and movements of the robot, the breaker, traffic, and background noise). The 20 dB attenuation level for the FS condition was defined during pilot studies when the participants reported it provided comfortable and informative levels of sound feedback during the simulations compared to the US condition.

The selection of the attenuated sound sources in the FS condition was made during the development of the virtual environment and subsequent pilot studies. For the sound sources associated with the robot's operational status (i.e., engine and breaker) and motions (i.e., tracks, arm, and outriggers movements), because the audio listener was positioned on the robot, these sounds were much louder than all other sounds during the simulations and, thus, were attenuated. The remaining attenuated sound sources (i.e., traffic and background noise) were attenuated because they were considered not to contribute to operations or inform about hazards and task conditions. The unattenuated sound sources (i.e., collapses, workers' voices, and equipment) remained unattenuated because they were relatively lower than the operational sounds of the robot and were regarded as informative of hazards and task conditions on-site. It is important to note that the categorization for attenuated and unattenuated sounds and the appropriate levels of sound attenuation may change for different applications, sites, or task conditions, and interface designers must identify proper combinations of attenuated/unattenuated sounds and attenuation levels for other applications.

In both experimental conditions, participants demolished two walls and two suspended slabs in another room of the same virtual building used in the baseline condition. Physiological measures (EDA and HR) were obtained throughout each condition, and SA, anxiety, and workload were measured after each experimental condition. Although the participants repeated the same demolition scenario in each condition, some differences existed between the two conditions to minimize learning effects in the assessment of SA. For that, the colors and materials of some built elements (e.g., walls, ceilings, slabs) were changed, the number of equipment and tools located outside the demolition area was altered, and the hazards related to collapse and agents entering the environment were modified. During the first experimental trial, regardless of the sound condition, the collapsing element was a wall, and the entering agent was another Brokk robot. During the second experimental trial, the collapsing element was a slab, and the entering agent was a worker.

4.5. Participants

The experiment was powered for small effect sizes (Cohen's d = 0.15). For two groups (order of applying the two conditions with sound included, i.e., US then FS or FS then US) and three measurements (i.e., NS (baseline), US, and FS), GPower suggested a sample size of 74 participants to achieve a power of 80 % with a $p < 0.05$ significance level. A convenience sample of eighty participants, all 18 years or older and fluent in English, was recruited through e-mail and the USC psychology subject pool. Participation was voluntary, and no compensation was provided to the participants. However, academic credits were granted to the participants recruited from the USC psychology subject pool. Five participants were dropped during the analysis of the results because three participants asked the researcher to reduce the sound levels of the headphones, and two participants skipped the baseline condition (NS). Table 1 presents the demographic information of the participants of the experiment by randomized experimental order sequence. Group A participants experienced the US condition followed by the FS condition, whereas Group B participants experienced the FS condition followed by the US condition.

4.6. Measurements and feature extraction

In this study, a variety of subjective and objective measures were used to assess the effects of environmental sound attenuation on the

Table 1
Demographic information of the participants (n = 80) by experimental condition order.

		Group A (US-FS)	Group B (FS-US)
Gender	Female	19 (24 %)	20 (25 %)
	Male	21 (26 %)	20 (25 %)
Age (years)	18–23	24 (30 %)	23 (29 %)
	24–29	12 (15 %)	13 (16 %)
	30 or older	3 (4 %)	4 (5 %)
Race/Ethnicity	African American	1 (1 %)	1 (1 %)
	Asian	18 (22 %)	22 (27 %)
	Caucasian	13 (16 %)	10 (12 %)
	Hispanic/Latino	4 (5 %)	5 (6 %)
	Middle Eastern	4 (5 %)	4 (5 %)
Completed Education	High School Degree /	23 (29 %)	20 (25 %)
	Some College	2 (3 %)	7 (9 %)
	Bachelor's Degree	15 (19 %)	13 (16 %)
	Master's degree		
Video Game Habits	Don't Play	14 (18 %)	22 (28 %)
	Play at Least Once a	6 (8 %)	6 (8 %)
	Year	5 (6 %)	4 (5 %)
	Play at Least Once a	8 (10 %)	5 (6 %)
	Month	7 (9 %)	3 (4 %)
	Play at Least Once a		
	Week		
	Play at Least Once a		
	Day		
Virtual Reality Experience	None	17 (21 %)	22 (28 %)
	Less than 5 Times	16 (20 %)	12 (15 %)
	5–15 Times	4 (5 %)	1 (1 %)
	More than 15 Times	3 (4 %)	5 (6 %)
Construction Experience	None	36 (45 %)	33 (41 %)
	Less than 5 Years	3 (4 %)	6 (8 %)
	5–10 Years	1 (1 %)	1 (1 %)
Construction Machine Operation Experience	None	38 (48 %)	39 (49 %)
	Less than 5 Times	2 (3 %)	1 (1 %)
Construction Machine Teleoperation Experience	None	40 (50 %)	40 (50 %)

Note: Participants were allowed to choose not to reply to any question and could select multiple options for some questions; therefore, not all variables total n = 80.

participants' task performance, safety performance, and response (i.e., experienced levels of workload, SA, anxiety, stress, attention, and task engagement) during teleoperated demolition.

4.6.1. Participants' task and safety performance

4.6.1.1. Average percent demolition. The average percent demolition (by volume) was computed for all elements to be demolished in each task scenario. In the NS condition, it considered one wall and one suspended slab. In the US and FS conditions, it considered two walls and two suspended slabs. Given the time constraints of the tasks and to prevent the participants from getting stuck on minor pieces of the target elements, a 70 % demolition threshold was used for each element (defined in a pilot study) to inform the participants to proceed to the next target.

4.6.1.2. Number of collisions. The number of collisions was computed as the count of collisions between the robot's body and other elements in the scenes (e.g., walls, columns, clutter). It did not include the collisions between the breaker and these elements during demolition. Also, to reduce the computational complexity of indicating different types and levels of collisions, no sound was included in the virtual environment to indicate the occurrence of collisions between the robot's body parts and objects in the environment.

4.6.1.3. Reaction time - collapse. The reaction time associated with collapsing elements was measured as the time the participant took to turn off the engine after an element collapsed. During the training session, participants were instructed to stop the engine to assess the situation if they perceived a collapse. No element in the demolition area collapsed during the NS condition to reduce the learning effects of the participants expecting an element to collapse in subsequent tasks. Alternatively, a wall collapsed during the second task, and a suspended slab collapsed during the third task. The maximum time a participant could react to these collapsing elements was 120 s, with the cracking and collapsing sounds provided during the first 20 s of this window.

4.6.1.4. Reaction time – agent. The reaction time associated with entering agents was measured as the time the participant took to turn off the robot's engine after the agents entered the demolition area. In the safety training, participants learned about the maximum reach of the Brokk's arm and were instructed to turn off the robot's engine if they perceived other equipment or workers approaching the robot. Two entering agents were used: a Brokk robot in the second task and a worker in the third task. Again, these agents were not used in the NS task to reduce learning effects. If the participants perceived the entering agents and reacted to them, the agents would leave the demolition area immediately. If the participants did not perceive the agents, the agents would leave the demolition area after 120 s. During the entire stay of these agents, the participant could hear the sounds from the entering Brokk or worker. For the Brokk, the sounds were from the engine and motions of the robot, and for the worker, the sounds were from the worker shouting instructions to another worker.

4.6.2. Participants' response

4.6.2.1. Workload. The subjective assessment NASA-TLX [43] was used to measure workload. This instrument is considered the gold standard for subjective workload assessments and has been shown to be reliable and valid [46]. In this experiment, the NASA-TLX was applied at the end of each of the three demolition tasks (NS, US, and FS). Each time, the average of the six subscales of the NASA-TLX instrument (i.e., mental demand, physical demand, temporal demand, performance, effort, and frustration) was calculated to determine the participant's workload.

4.6.2.2. Situation Awareness (SA). The Situation Awareness Global Assessment Technique (SAGAT) [47,48] was used to measure the participants' SA. We created a pool of 23 questions related to the site and operating conditions of the robot to assess the overall awareness of the participants about the environment, task, and existence of hazards on site. After each experimental condition, the participants responded to six randomly selected questions, two from each of the three levels of SA: level 1 - perception of the environment, level 2 – comprehension of its meaning, and level 3 – projection of the near future. Although no questions were repeated following one condition, the same questions could be repeated following both conditions, which is why the environment's components between the two conditions were altered. Then, the number of correct answers was used to determine the SA scores of each participant at each task condition.

4.6.2.3. Anxiety. The 6-item version of the Spielberger State-Trait Anxiety Inventory (STAI) [42] was used to measure anxiety. This instrument is highly correlated with the longer 20-item STAI version, and all internal consistency reliabilities are greater than 0.90 [49]. Trait anxiety refers to an individual's predisposition to experience anxiety in stressful situations, while state anxiety refers to a temporary emotional response to a specific event [50]. The 6-item STAI instrument was used five times during the experiment: before the training session, after the participants' physiological baseline assessments, and once after each of the three demolition tasks.

Integers were assigned to the Likert scale options for each instrument item to compute the STAI scores of the participants. For the "anxiety-present" items (i.e., tense, upset, and worried), the assignments were: Not at all = 1, Somewhat = 2, Moderately so = 3, Very much so = 4. Alternatively, for the "anxiety-absent" items (i.e., calm, relaxed, and content), the assignments were: Not at all = 4, Somewhat = 3, Moderately so = 2, Very much so = 1. The values of the six items were summed, resulting in a number between 6 and 24, which was multiplied by the ratio 80/24 to scale the total score to the same range as the 20-item STAI instrument.

4.6.2.4. Stress. To predict the development of stress in the participants, change scores relative to baseline measurements and absolute scores of a variety of HR and heart rate variability (HRV) metrics were used. Prior research shows that stress affects HRV and that HRV can be used in objective assessments of stress [51]. A Polar H10 sensor [52] was used, and over 40 HRV parameters were extracted using the Kubios HRV Scientific software [53]. Kubios HRV was used for data pre-processing, feature extraction, and segmentation. During data pre-processing, a visual inspection of the quality of the ECG data for each participant was performed. Additionally, the software triggered low-quality warnings when more than 5 % of the data needed beat correction, which supported quality inspection. After quality inspection, a medium-level automatic noise detection filter was applied to the data, which excluded noise segments from subsequent analysis. Alternatively, for individual intermittent abnormal beat intervals, Kubios HRV corrected them automatically using a beat correction algorithm. For data segmentation and feature extraction, the events marked on the Empatica E4 wristband were used to find the beginning and end of each demolition task. For the analysis, we focused on the mean RR interval, mean HR, and some of the most reported HRV metrics: the standard deviation of RR intervals (SDNN), the root mean square of successive differences (RMSSD), and the ratio between low-frequency and high-frequency band powers (LF/HF) [54].

The mean RR intervals refer to the mean values of the intervals between successive R-waves in the electrocardiogram (ECG) signal. Using the values of the RR intervals from the Polar H10, many time-domain features were extracted, including HR (i.e., the number of heartbeats per minute), SDNN, and RMSSD. We also extracted the LF/HF ratio, which is part of the frequency-domain features from the signal, and it is commonly used to estimate the ratio between the sympathetic nervous system (SNS) activity and the parasympathetic nervous system (PNS) activity [55]. SNS and PNS are the two divisions of the autonomic nervous system (ANS), with the PNS being predominant in "rest and digest" situations and the SNS in more stress-inducing "fight or flight" situations [56]. For all HRV indices, except the LF/HF ratio, changes in the HRV features relative to baseline measurements were calculated. Using change scores for each participant reduces the effects of personal differences, allowing for a more reliable between-subject evaluation of the effect of each task condition on physiological responses. The reason actual LF/HF ratios were used instead of changes in LF/HF ratios from the baseline measurements was that slow respiration rates during resting periods significantly influence the LF band power, which affects the interpretation of baseline LF/HF ratios [55].

4.6.2.5. Attention and task engagement. To predict the participants' attention and task engagement levels during the demolition tasks, change scores relative to baseline measurements and absolute scores of various EDA features were used. Prior research has shown a relationship between Skin Conductance Responses (SCR) and participants' levels of attention and task engagement [57,58]. The study of EDA focuses on the autonomic changes in the electrical properties of the skin, such as skin conductance [59]. EDA has two components: (1) tonic-level EDA, measured in terms of Skin Conductance Level (SCL), refers to background or slow-change skin conductance signals, and (2) phasic-level EDA, measured in terms of SCR, refers to fast-change skin conductance signals that can be attributable to specific stimuli in some cases [59].

The Empatica E4 wristband [60] was used to measure the participants' EDA. EDA data was processed in the biosignal-specific processing toolbox (Bio-SP tool, version 2.2) [61] to mitigate the effects of noise on the data and extract various features related to the SCRs. First, the data was segmented based on the events marked on Empatica E4, i.e., the beginning and end of each demolition task. Then, a Gaussian low-pass filter with a window of 40 points and a sigma of 400 ms was used to reduce the effects of noise in the EDA data, as presented in [62]. Finally, the SCR features were extracted from the filtered data. These features included the SCR duration mean, SCR amplitude mean, SCR rise-time mean, and the number of detected SCRs. For the subsequent analysis of the results, except for the number of SCRs, the change scores in the EDA features relative to baseline measurements were considered.

4.7. Data analysis

For each variable in the study, the descriptive statistics were computed, and the normality of residuals and the existence of outliers were assessed for each group and condition. For normality assessments, the Kolmogorov-Smirnov test was used. For outlier detection, the data's interquartile range (IQR), defined as the difference between the third and first quartiles, was computed for each group and condition. Then, the upper and lower limits of the data were determined by adding three times the IQR to the third quartile and subtracting three times the IQR from the first quartile, respectively. Any values outside this interval (i.e., outliers) were removed from subsequent analyses. Then, the Kolmogorov-Smirnov test was applied again for the data with outliers removed. In a few cases, some groups and conditions still did not meet the normality of residuals assumption required by mixed ANOVAs. However, even in these cases, mixed ANOVAs were used because they have been shown to be robust to violations of normality assumptions for sufficiently large sample sizes [63,64]. During data analysis, deviations from the sphericity assumption were assessed using Mauchly's test of sphericity. When violations of the sphericity assumption were observed, the degrees of freedom were corrected using the Greenhouse–Geisser correction for situations where $\varepsilon < 0.75$ or the Huynh–Feldt correction for situations where $\varepsilon > 0.75$.

Mixed ANOVAs were used to examine differences by condition (i.e., NS (baseline), FS, and US) and experimental order (i.e., Group A (US-FS) and Group B (FS-US)). The mixed ANOVA analysis varied by the available data for each outcome. For example, reaction times and SA were not evaluated in the NS condition and, thus, were only examined between the FS and US conditions. For each variable, we first analyzed whether a significant order interaction existed; if that was the case, simple main effects were reported. If no significant interaction was observed, the main effects were analyzed.

For pairwise comparisons for simple main effects and main effects, Fisher's least significant difference (LSD) was used. The decision to use Fisher's LSD test instead of a more conservative test such as Bonferroni's followed the study's objectives of minimizing the chance of missing any effects during pairwise comparisons (type II error), as Bonferroni's test can be unnecessarily conservative in these cases [65]. All statistical analyses were conducted using IBM's Statistical Package for the Social Sciences (SPSS) (version 28).

5. Results

Table 2 presents the descriptive statistics (means, standard deviations, and sample sizes) for each variable of interest in each sound condition and group. Comparative analysis of measures among the conditions are grouped and presented in two sections related to (1) task and safety performance (i.e., percent demolition, collisions, and reaction times to collapses and agents) and (2) participants' psychological and physiological responses (i.e., SA, workload, and anxiety as measured subjectively, stress as predicted by HRV features, and attention and task engagement as predicted by the SCR from the EDA signal).

5.1. Participants' task and safety performance

Fig. 3 displays the estimated marginal means, standard errors, and significant differences between sound conditions for the four performance measures of average percent demolition (Figs. 3 (a) and 3 (b)), number of collisions (Fig. 3 (c)), reaction time to collapses (Fig. 3 (d)), and reaction time to entering agents (Fig. 3 (e)).

5.1.1. Average percent demolition

No interaction effect between group and sound condition was observed for average percent demolition ($p = 0.203$) using a 2×3 mixed ANOVA with a Huynh-Feldt correction of the degrees of freedom ($\varepsilon = 0.835$). However, a significant effect of sound condition was found ($F(1.669, 105.164) = 17.936$, $p < 0.001$, $\eta_p^2 = 0.222$) due to significant differences between the NS condition and each of the two experimental conditions ($p < 0.001$ in both cases) (Fig. 3 (a)). These differences may demonstrate improved performance in FS and US conditions compared to NS, or they may result from a learning effect since the NS condition was the first task in groups A and B. Alternatively, no significant difference in percent demolition was found between the FS and US conditions ($p = 0.198$). Finally, a significant effect of group for the percent demolition was observed ($F(1, 63) = 9.282$, $p = 0.003$, $\eta_p^2 = 0.128$) (Fig. 3 (b)) with the percent demolition being higher for group A than for group B. However, we could not find an explanation for this observation.

We used 2×2 mixed ANOVAs to compare the two experimental sound conditions for the other three performance measures, i.e., number of collisions, reaction time to an agent, and reaction time to a collapse.

5.1.2. Number of collisions

The number of collisions (Fig. 3 (c)) had no interaction effect ($p = 0.056$) and no main effect by group ($p = 0.480$); however, a main effect of sound condition was detected ($F(1, 63) = 9.648$, $p = 0.003$, $\eta_p^2 = 0.133$) as significantly more collisions occurred in the US condition than in the FS condition.

5.1.3. Reaction time – collapse

For the reaction times to collapsing elements (Fig. 3 (d)), a main

Table 2
Descriptive statistics (means, standard deviations, and sample sizes) for the measures of interest.

Measure[‡]		Sound Condition									Group	
		No Sound Condition			Filtered Sound Condition			Unfiltered Sound Condition				
		Group A[*]	Group B[†]	Total	Group A	Group B	Total	Group A	Group B	Total	A	B
Participants' Task and Safety Performance												
Average Demolition (%)	M	62.2	58.1	**60.0**	78.0	65.2	**71.3**	77.3	70.5	**73.7**	**72.5**	**64.6**
	SD	18.3	21.8	**20.2**	5.1	17.4	**14.5**	7.2	14.8	**12.2**	**1.9**	**1.8**
	N	31	34	**65**	31	34	**65**	31	34	**65**	**31**	**34**
Number of Collisions	M	–	–	–	10.5	12.0	**11.3**	18.2	13.8	**15.9**	14.4	12.9
	SD	–	–	–	5.5	5.3	**5.4**	17.3	8.8	**13.7**	1.5	1.5
	N	–	–	–	32	33	**65**	32	33	**65**	32	33
Reaction Time – Collapse (s)	M	–	–	–	53.1	57.9	**55.6**	87.4	65.9	**76.4**	70.3	61.9
	SD	–	–	–	54.8	53.7	**53.8**	48.4	56.7	**53.5**	7.9	7.6
	N	–	–	–	33	35	**68**	33	35	**68**	33	35
Reaction Time – Agent (s)	M	–	–	–	**16.1**	**85.0**	58.4	**113.8**	**50.2**	74.7	64.9	67.6
	SD	–	–	–	**17.2**	**44.6**	49.6	**18.9**	**42.9**	47.2	5.9	4.7
	N	–	–	–	**22**	**35**	57	**22**	**35**	57	22	35
Participants' Response												
Workload	M	37.1	41.4	39.3	35.0	41.1	38.1	38.5	37.9	38.2	36.9	40.1
	SD	15.0	18.7	17.0	15.8	18.6	17.4	17.1	18.5	17.7	2.5	2.5
	N	37	38	75	37	38	75	37	38	75	37	38
Situation Awareness (SA)	M	–	–	–	4.2	3.8	4.0	3.8	4.1	4.0	4.0	4.0
	SD	–	–	–	1.1	1.2	1.2	1.3	1.1	1.2	0.2	0.1
	N	–	–	–	33	35	68	33	35	68	33	35
State-Trait Anxiety Inventory (STAI)	M	37.1	37.2	37.2	37.3	36.7	37.0	38.7	36.2	37.5	37.7	36.7
	SD	13.0	12.0	12.4	11.8	10.9	11.3	12.5	11.7	12.1	1.8	1.8
	N	37	38	75	37	38	75	37	38	75	37	38
Δ Mean RR (ms)[§]	M	−11.2	−10.1	−10.7	−3.0	−8.7	−5.9	−8.8	−13.4	−11.2	−7.7	−10.8
	SD	46.2	39.1	42.4	51.6	35.5	43.9	56.5	38.9	47.9	7.0	6.8
	N	36	38	74	36	38	74	36	38	74	36	38
Δ Mean Heart Rate (HR) (beats/min)[§]	M	1.4	0.7	1.1	0.9	0.7	0.8	1.0	1.3	1.1	1.1	0.9
	SD	4.9	3.1	4.1	5.1	3.3	4.2	5.8	4.0	4.9	0.7	0.7
	N	35	38	73	35	38	73	35	38	73	35	38
Δ SDNN (ms)[§]	M	−9.6	−9.2	**−9.4**	−7.0	−6.8	**−6.9**	−7.9	−6.5	**−7.2**	−8.2	−7.5
	SD	8.0	10.6	**9.4**	6.9	10.9	**9.2**	8.5	10.6	**9.6**	1.5	1.5
	N	34	36	**70**	34	36	**70**	34	36	**70**	34	36
Δ RMSSD (ms)[§]	M	−5.4	−1.4	−3.4	−4.2	−0.4	−2.3	−4.8	−1.8	−3.3	−4.8	−1.2
	SD	8.2	8.6	8.6	8.6	7.9	8.4	9.2	9.2	9.2	1.4	1.4
	N	35	35	70	35	35	70	35	35	70	35	35
LF/HF Power Ratio[§]	M	2.0	2.0	**2.0**	1.8	2.4	**2.1**	2.3	2.9	**2.6**	2.1	2.4
	SD	1.3	1.5	**1.4**	1.0	1.8	**1.5**	1.6	2.3	**2.0**	0.3	0.3
	N	33	36	**69**	33	36	**69**	33	36	**69**	33	36
Δ SCR Duration Mean (ms)[¶]	M	−7.4	−4.3	−5.8	−8.1	−4.7	−6.3	−10.3	−3.4	−6.8	−8.6	−4.1
	SD	17.8	12.8	15.4	21	12.7	17.2	19.5	13.7	17	2.7	2.6
	N	34	36	70	34	36	70	34	36	70	34	36
Δ SCR Amplitude Mean (μS)[¶]	M	0.08	0.06	**0.07**	0.20	0.11	**0.15**	0.05	−0.01	**0.02**	0.11	0.05
	SD	0.21	0.19	**0.20**	0.53	0.26	**0.41**	0.30	0.13	**0.23**	0.04	0.04
	N	25	27	**52**	25	27	**52**	25	27	**52**	25	27
Δ SCR Rise-Time Mean (ms)[¶]	M	−5.2	0.5	−2.3	−4.4	−0.2	−2.3	−6.1	−0.2	−3.1	**−5.2**	**0.0**
	SD	10.7	6.0	9.0	12.8	5.5	10.0	10.3	7.1	9.2	**1.5**	**1.4**
	N	35	36	71	35	36	71	35	36	71	**35**	**36**
Number of SCRs[¶]	M	24.1	24.2	24.1	22.8	22.7	22.8	24.2	25.8	25.0	23.7	24.2
	SD	13.6	15.9	14.7	13.4	16.3	14.8	14.0	15.0	14.4	1.8	1.8
	N	37	38	75	37	38	75	37	38	75	37	38

Note: Highlighted results (in bold) indicate significant differences ($p < 0.05$). See Sections 5.1 and 5.2.
Abbreviations for the units: seconds (s), milliseconds (ms), minutes (min), micro siemens (μS).
‡ Mean (M), standard deviation (SD), and sample size (N). Differences in sample sizes by group and condition result from missing data and removed outliers.
* Group A order of tasks: NS – US – FS.
† Group B order of tasks: NS – FS – US.
§ HR and HRV metrics used as predictors of stress.
¶ EDA metrics used as predictors of attention and task engagement.

effect of sound condition was observed ($F(1, 66) = 9.273$, $p = 0.003$, $\eta_p^2 = 0.123$), with the reaction times being faster in the FS condition than in the US condition. No interaction effect between sound condition and group ($p = 0.062$) or main effect of group were observed ($p = 0.449$).

5.1.4. *Reaction time – agent*

Alternatively, for the reaction times to entering agents, an interaction effect between group and sound condition was found ($F(1, 55) = 111.345$, $p < 0.001$, $\eta_p^2 = 0.669$). Reaction times to the entering agent (Fig. 3 (e)) decreased significantly between the first and second experimental conditions in each group ($p < 0.001$ in both cases), which may suggest that participants improved their ability to perceive entering agents as the experiment progressed. However, the observed decreases were larger for group A than for group B. In group A, reaction times decreased from 113.8 s (US) to 16.1 s (FS), while for group B, reaction times decreased from 85.0 s (FS) to 50.2 s (US). For the reaction times to an entering agent, we also observed a significant main effect of sound condition ($F(1, 55) = 25.081$, $p < 0.001$, $\eta_p^2 = 0.313$), but no main effect of group ($p = 0.726$).

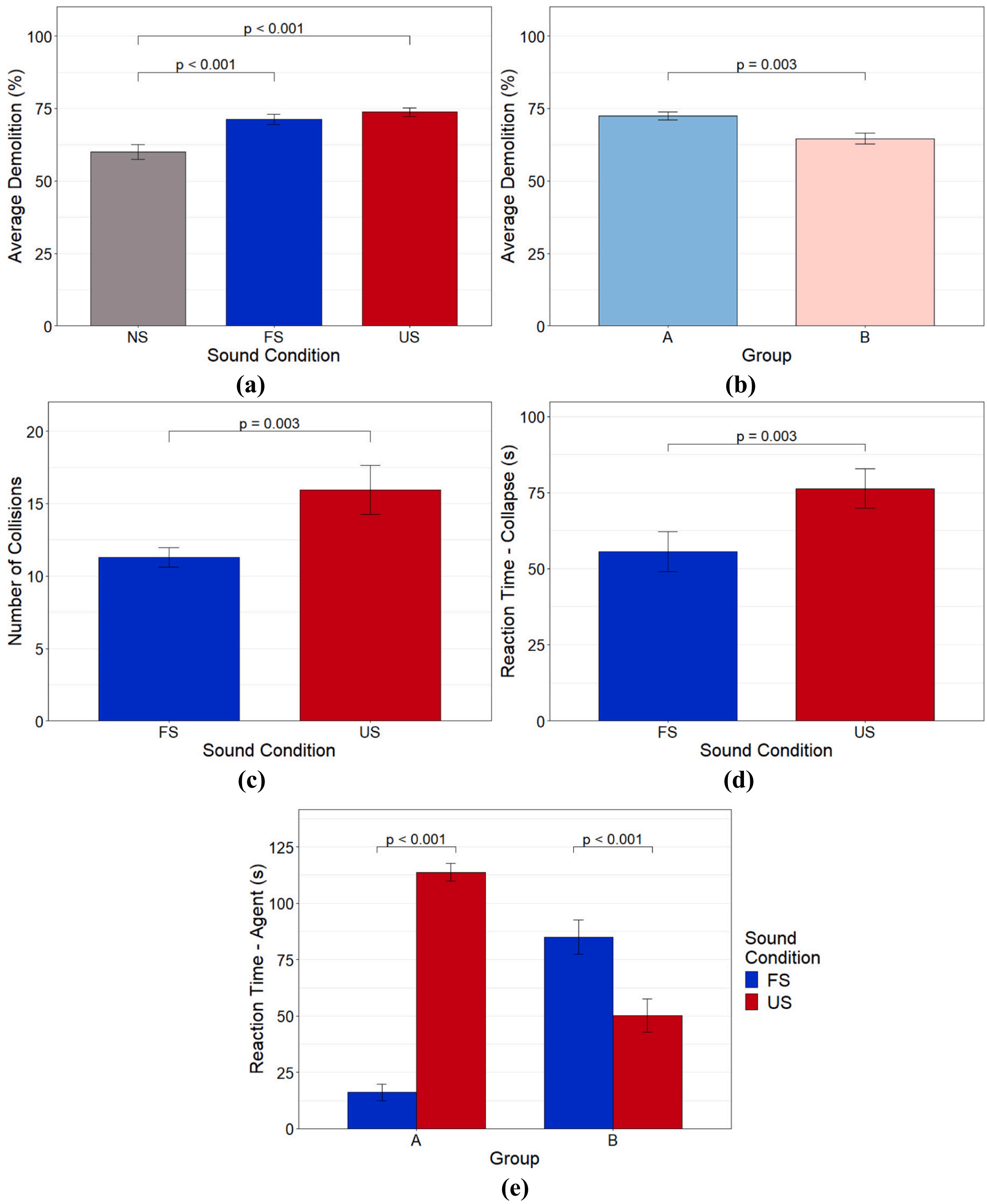


Fig. 3. Comparison of estimated marginal means between (a) sound conditions for average demolition showing higher demolition performance in FS and US compared to NS, (b) groups for average demolition showing higher demolition performance in group A than in group B, (c) number of collisions showing fewer collisions in FS than US, and (d) reaction time to collapses showing faster reaction times in FS than US, and (e) between conditions by participant group for reaction time to agents showing faster reactions times in the second experimental conditions than in the first experimental conditions in each group. Error bars indicate standard errors of the mean.

5.2. Participants' response

5.2.1. Workload

Workload was examined using a 2 × 3 mixed ANOVA with a Huynh-Feldt correction of the degrees of freedom ($\varepsilon = 0.899$). No significant interaction effect between sound condition and group ($p = 0.137$) or significant main effect by sound condition ($p = 0.708$) or group ($p = 0.358$) were observed for workload.

5.2.2. Situation awareness (SA)

Similarly, in the analysis of SA using a 2 × 2 mixed ANOVA, no interaction effect between sound condition and group ($p = 0.063$) or main effects of sound condition ($p = 0.832$) or group ($p = 0.785$) were observed.

5.2.3. Anxiety

Anxiety differences were examined using a 2 × 3 mixed ANOVA. No interaction effect between sound condition and group ($p = 0.305$), and no main effects of sound condition ($p = 0.841$) or group ($p = 0.694$) were observed. The STAI scores for all three sound conditions stayed below the threshold of at least 40, which is used to detect clinical levels of anxiety [66].

5.2.4. Stress

We used 2 × 3 mixed ANOVAs to examine the HR and HRV measures used as predictors of stress. No interaction effects between sound condition and group, and no main effects of group were found for any of the HRV measures reported in this study (all $p > 0.05$). There were no main effects of sound condition for changes in RR intervals, HR, and RMSSD (all $p > 0.05$). Significant differences in the main effects of sound condition were noted in changes in SDNN relative to baseline measurements and LF/HF ratio.

In the analysis of changes in SDNN, no interaction between sound condition and group ($p = 0.577$) and no main effect of group ($p = 0.749$) were observed, but a significant main effect of sound condition was found ($F(2, 136) = 9.553$, $p < 0.001$, $\eta_p^2 = 0.123$). Pairwise comparisons of SDNN showed significant differences between the NS and FS conditions ($p < 0.001$) and the NS and US conditions ($p = 0.001$) but no difference between the FS and US conditions ($p = 0.628$) (Fig. 4 (a)). SDNN decreased from baseline measurements in all three conditions, with the NS condition showing the biggest decrease. The reduction in the magnitude of the decrease from the first demolition task (NS) to subsequent tasks (US and FS), however, may suggest habituation to the demolition tasks or decrements in basal emotion levels as the experiment progressed [67]. SDNN can generally be considered an indicator of physiological resilience to stress [51]. Existing results in the literature show that SDNN increases during relaxation periods due to the recovery of parasympathetic activity [68] and decreases during stress periods [69]. Thus, the results may suggest increased stress levels during the three demolition tasks relative to the baseline measurements, regardless of sound condition.

For the LF/HF power ratio in the Fast Fourier Transform (FFT) spectrum, no interaction effect between sound condition and group ($p = 0.060$) or main effect of group ($p = 0.311$) was observed. However, a significant main effect of sound condition ($F(2, 134) = 9.733$, $p < 0.001$, $\eta_p^2 = 0.127$) was found. Pairwise comparisons showed significant differences between the US and NS conditions and the US and FS conditions ($p < 0.001$ in both cases) but no significant difference between the NS and FS conditions ($p = 0.541$) (Fig. 4 (b)). The US condition had the highest LF/HF ratio among the three sound conditions. As presented, higher LF/HF ratios can be associated with higher levels of sympathetic dominance [55], which is more common in stress-inducing "fight or flight" situations. Thus, the results may indicate that participants experienced higher stress levels in the US condition than in the NS and FS conditions. Fig. 4 presents the estimated marginal means and standard errors of the (a) change in SDNN relative to the baseline measurements and (b) LF/HF power ratio for all sound conditions.

5.2.5. Attention and task engagement

2 × 3 mixed ANOVAs were used to examine the EDA measures selected as predictors of attention and task engagement. No interaction effects between sound condition and group were found for any EDA measures reported in this study (all $p > 0.05$). There were no main effects of sound condition for changes in SCR duration means, SCR rise-time means, and number of SCRs, and no main effects of group for changes in SCR duration means, SCR amplitude means, and number of SCRs (all $p > 0.05$). Significant differences were observed in the main effect of sound condition in SCR amplitude means and in the main effect of group in SCR rise-time means.

In our analysis of the SCR amplitude means with a Huynh-Feldt correction of the degrees of freedom ($\varepsilon = 0.814$), no significant interaction between sound condition and group ($p = 0.741$) and no main effect of group ($p = 0.351$) were observed, but a significant main effect of sound condition was found ($F(1.628, 81.410) = 4.121$, $p = 0.027$, $\eta_p^2 = 0.076$). Pairwise comparisons showed significant differences between FS and US ($p = 0.011$) but no significant differences between NS and US ($p = 0.153$) or NS and FS ($p = 0.126$) (Fig. 5 (a)). SCR amplitude means increased for all sound conditions relative to baseline measurements, with the FS condition having a higher increase than the US and NS conditions. Existing results in the literature show that SCR amplitudes relate to participants' attention and task engagement levels [57,58], suggesting that the FS condition induced the highest levels of attention and task engagement among the three sound conditions.

Finally, for the SCR rise-time means with a Huynh-Feldt correction of the degrees of freedom ($\varepsilon = 0.925$), no significant interaction between sound condition and group ($p = 0.313$) and no main effect of sound condition ($p = 0.303$) were found. However, a significant main effect of group was observed ($F(1, 69) = 6.704$, $p = 0.012$, $\eta_p^2 = 0.089$) (Fig. 5 (b)). The SCR rise-time mean is the mean time between the start of the SCR events and their respective peak values. Group A had a mean decrease of 5.24 ms in SCR rise-time means from baseline

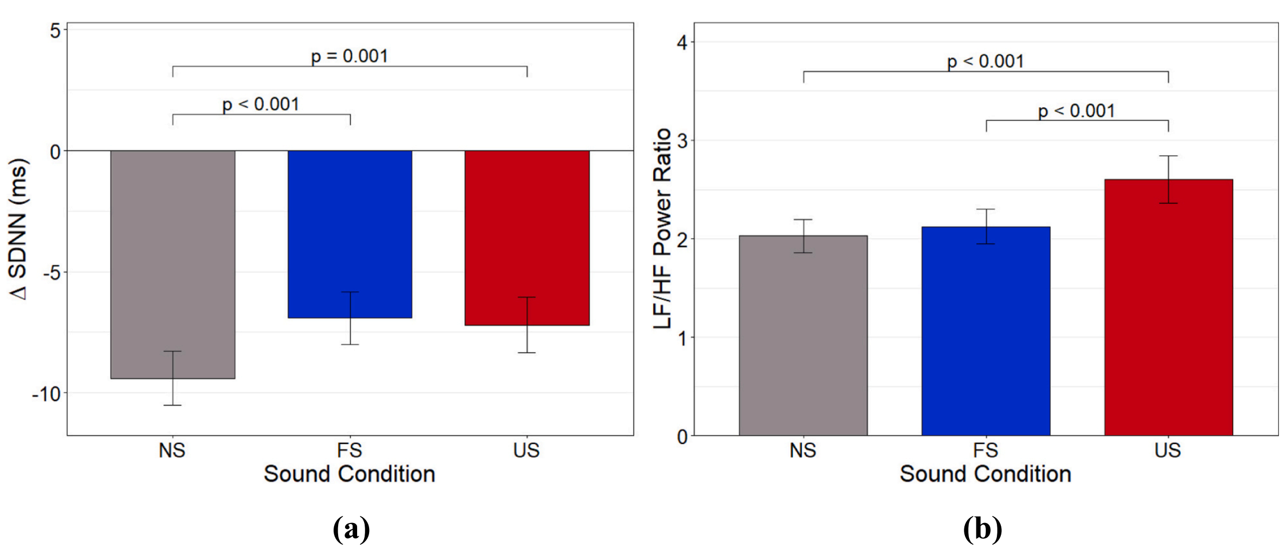


Fig. 4. Comparison of estimated marginal means between sound conditions for the Heart Rate Variability (HRV) parameters (a) changes in SDNN relative to baseline measurements likely showing increased levels of stress associated with the demolition tasks with habituation to the task over time and (b) LF/HF power ratio (FFT spectrum) likely showing higher levels of stress in US than in NS and FS conditions. Error bars indicate standard errors of the mean.

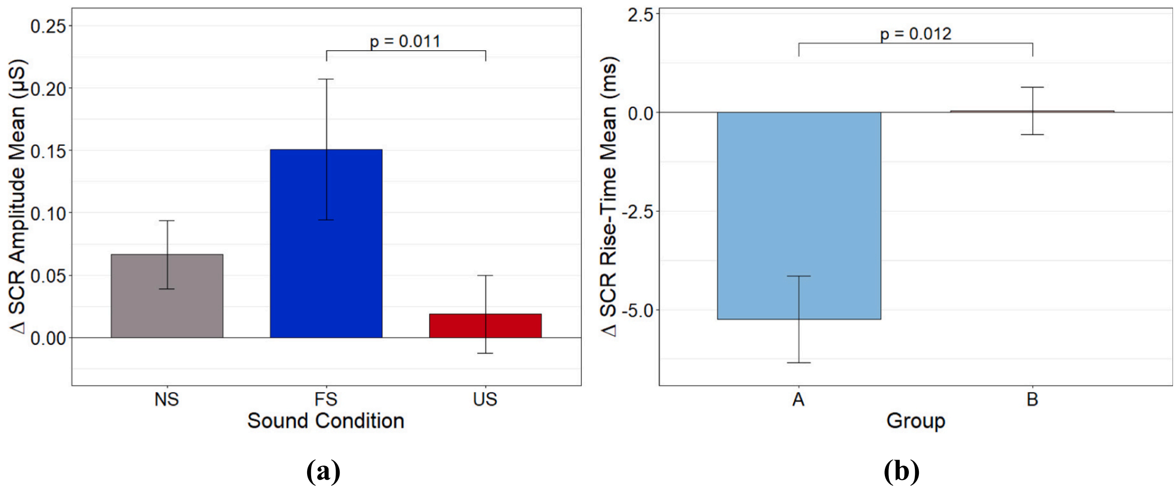


Fig. 5. Comparison of estimated marginal means for changes in Electrodermal Activity (EDA) features (a) between sound conditions for changes in SCR amplitude means likely showing higher levels of attention and task engagement in FS than US and (b) between groups for changes in SCR rise-time means relative to baseline measurements of unclear significance. Error bars indicate standard errors of the mean.

measurements, while group B had a mean increase of 0.04 ms. Decreases in SCR rise-time have also been associated with increased attention levels [70]. Thus, the results may indicate higher attention levels in group A than in group B; however, we could not find an explanation for this observation. Fig. 5 presents the estimated marginal means and standard errors of the change in (a) SCR amplitude means for all sound conditions and (b) SCR rise-time means for all groups relative to baseline measurements.

6. Interpretation of results and discussion

6.1. Effects on operators' task and safety performance

In answering research question 1, the study's results show that, compared to unfiltered sound (US condition), selective environmental sound attenuation (FS condition) improved operators' safety performance but did not significantly affect their task performance. Specifically, the FS condition resulted in fewer collisions and faster reaction times to hazards compared to the US condition (Figs. 3 (c) – (e)). For the number of collisions, the higher sound levels in the US condition likely impacted the participants' driving and positioning performance, thus resulting in more collisions than in the FS condition. Prior research has shown that the levels of acoustic support (i.e., type and amount of auditory information) affect positioning accuracy in teleoperation [20]. Therefore, the focused sound conditions in FS may have contributed to improved driving and positioning performance, potentially explaining fewer collisions in FS than in US. For the differences in reaction times to hazards, the attenuated sounds for the robot and background noise in FS might have facilitated the perception of the entering agents and collapsing elements, potentially resulting in faster reaction times in FS than in US. Specifically, the results showed faster reaction times in the US condition compared to the FS condition for the reaction times to collapsing elements. Alternatively, the results showed a potential learning effect for the reaction times to entering agents, as participants reacted faster to entering agents during the second experimental condition than in the first experimental condition in each group. However, the results also showed that the reduction in reaction times was more pronounced when participants experienced the FS condition after the US condition (Group A) than when they experienced the US condition after the FS condition (Group B). Previous research has shown that including environmental sounds in virtual environments can help users perceive certain hazards more easily [28]. Thus, in the FS condition, the fact that it was easier to perceive certain sounds associated with hazards (e.g., collapsing elements, entering agents) due to the attenuation of the sounds associated with the robot's operational status and background noise may have made it easier to identify these hazards and react to them. Alternatively, no significant differences between the FS and US conditions were detected for task performance as measured by percent demolition and SA.

6.2. Effects on operators' response: stress, attention, and task engagement

Regarding research question 2, which focused on the effects of selective environmental sound attenuation (FS condition) on operators' response (e.g., stress, attention, task engagement) compared to unfiltered sound (US condition), the physiological indicators of stress, attention, and task engagement suggested improved operator responses in FS compared to US. From the results in Table 2, the increases in mean HR and the decreases in mean RR intervals, SDNN, and RMSSD relative to baseline physiological measurements may indicate increased levels of stress for the three tasks, regardless of sound condition, as increases in HR and decreases in heart rate variability have been commonly used as indicators of increased levels of stress [51]. For the case of SDNN (Fig. 4 (a)), specifically, the decrease was significantly greater in the NS condition than in the US and FS conditions, suggesting less stress resilience in NS than in US and FS, but since the NS condition was always the first condition the participants experienced, this observation may be related to habituation to the demolition task [67]. The fact that increased stress levels were observed for all three tasks relative to baseline measurements indicates that the simulated demolition tasks triggered a physiological response in the participants. However, these measures alone did not let us completely characterize the contributions of the different sound conditions on the observed increases in stress levels. To that end, the LF/HF results provided additional information on the effects of the sound conditions on the potential stress levels the participants experienced during the demolition tasks. As Fig. 4 (b) shows, the LF/HF ratio was significantly higher in the US condition than in the NS and FS conditions. Higher LF/HF ratios are commonly associated with higher levels of sympathetic dominance in stress-inducing "fight or flight" situations [55]. Thus, the results may indicate higher stress levels in the US condition than in the NS and FS conditions, which can be associated with the relatively higher levels of sound and the unfocused nature of the sound feedback the participants experienced in the US condition.

Significant differences between the FS and US conditions were also observed for the EDA signals (Fig. 5 (a)). As presented, SCR amplitudes have been associated with the participants' levels of attention and task engagement [57,58]. Significant sound stimuli can orient attention and increase task engagement, while abrupt sounds elicit startling responses that can lead to task disengagement [58]. In the FS condition, the attenuated sound levels for the robot's engine, robot's motions, and background noise may have improved the participant's ability to identify specific events on-site compared to the unattenuated sounds in the US condition. For example, the faster reaction times in the FS condition suggest that the participants had a better capacity to discern these sounds during the demolition task in the FS condition than in the US condition. Thus, the focused sound conditions in the FS condition likely resulted in higher attention and task engagement levels, which may explain the higher SCR amplitude means in the FS condition compared to the US condition. Alternatively, subjective measures of workload and anxiety and some physiological variables did not indicate any significant differences in their effects among sound conditions, groups, or interaction effects between sound conditions and groups.

Finally, the failure to observe significant differences in demolition performance (Fig. 3 (a)) and SA between the FS and US conditions, if any exist, can be explained by existing stress-response models. These models state that task output is affected only when the adaptive capacities of the physiological and psychological responses to stress are exceeded [71] and that, for most tasks, it is possible to maintain similar levels of performance in the presence of a stressor by increasing the effort put into the task [72]. As the results show, no differences in demolition performance and SA between the FS and US conditions were observed, but the selected indicators of stress, attention, and task engagement suggested relatively worse physiological responses in US than in FS. Since higher sound levels are considered more uncomfortable than lower sound levels [17], and the sound feedback was less focused in the US condition, these results may suggest higher levels of physiological adaptation in US than in FS to achieve the observed levels of demolition performance in both sound conditions.

6.3. Further considerations

The results of this study make several contributions to the fields of teleoperated construction and human-factors engineering. As presented, there is an increasing interest in mechanisms that facilitate the perception of safety-critical sounds in construction sites [14], where noisy conditions still lead to accidents and occupational injuries [4,34]. While previous studies have focused on the importance of environmental and synthetic sounds for hazard perception in teleoperation scenarios, this study investigated the role of selective sound attenuation on operators' performance and response (e.g., stress, attention, task engagement). The study's findings show that selective environmental sound attenuation in teleoperation can help operators react faster to hazards and avoid collisions while experiencing higher levels of attention and task engagement and lower levels of stress. Thus, this paper highlights potential avenues to optimize sound feedback in teleoperation to improve operators' performance and response. It also bridges an important gap in understanding how noise affects cognitive and sensory processes in teleoperation, paving the way to developing enhanced teleoperation interfaces that prioritize safety-critical sounds.

Although the attenuated sound sources in the experiment were customized for demolition applications, the study's findings can still benefit other teleoperation applications in construction. By identifying the benefits of attenuating non-critical environmental sounds, the study highlights how technology can mitigate the inherent challenges posed by high-noise environments, leading to better safety outcomes and operators' responses. These insights can be applied to other areas of teleoperated work, such as disaster response, mining, and manufacturing, where clear and precise sound feedback is critical to safety and performance. However, to realize these benefits in other teleoperation applications, it is essential to identify the critical sounds relevant to their specific tasks and safety requirements.

Existing research gives indications on how to achieve the FS condition in practice. For example, ML models can be trained to detect anomalous sounds associated with hazards, accidents, and equipment malfunctions [35] or to enhance spoken communication among workers [38]. In spatial audio applications, existing algorithms can locate and separate audio sources, amplify desired sounds, and attenuate ambient noise [73]. Beyond the variables measured in this study, these attenuation mechanisms can potentially contribute to other performance indicators. Examples include increasing operators' comfort and improving their perception of the system's usability, helping them better recognize the robot's operational conditions and detect issues with it, and making it easier to understand spoken communications from on-site workers. However, these effects remain unexplored and present valuable opportunities for future research.

6.4. Limitations

Our findings offer valuable insights into the effects of selective environmental sound in teleoperated demolition and could inform the design of more effective teleoperation interfaces for demolition robots. However, some limitations also exist and could be addressed by future research efforts. First, all participants were students and not demolition robot operators; the participants' lack of professional demolition experience may affect the generalizability of the results. Nevertheless, the findings provide valuable insights into key factors that enhance operator performance and offer broader benefits for other teleoperated construction applications. Secondly, while virtual environments have been successful training tools [74], they are not real working environments. This study used a computer simulation instead of a real teleoperated demolition robot at a construction site, which was justified from a cost and safety perspective. Additionally, given the nature of the proposed intervention, using a simulation in a controlled environment allowed us to maintain the desired sound levels at each environmental sound condition, which would be far more complex to achieve in a real construction site. Therefore, differences in the observed effects of environmental sound attenuation might be found in validation studies conducted in real teleoperated demolition settings.

Limitations of the experiment included how chosen procedure characteristics (time intervals for tasks, sound levels chosen) and data collection practices may have affected outcomes. First, the overall duration of the experiment, about 1 h and 15 min in total, may have induced some level of fatigue in the participants, potentially affecting their performance and psychological and physiological responses as the experiment progressed. However, the participants' responses that could indicate increased levels of fatigue over time during the experiment (e.g., workload, HRV metrics) did not show this effect. Additionally, the experiment task durations, limited to 10 min per sound condition, may not have been sufficient to trigger stronger psychological and physiological responses in the participants. This duration is significantly shorter than the construction workers' typical 8 h/day work shifts.

Additionally, the fixed sound volumes for the US and FS conditions may have affected the participants' performance, stress, attention, and task engagement to some extent. While necessary for experimental control, this approach overlooked individual preferences for sound volumes. As presented, some participants asked the researcher to reduce the volume of their headsets due to discomfort, while others reported the sound levels to be comfortable. Finally, the movements of the participants' arms to control the joysticks may have introduced additional noise in the EDA data. For many studies involving the Empatica E4 sensor, participants are asked to wear the wristband in their nondominant arms to minimize motion noise; however, the control mechanisms of the simulated robot in this study did not allow for limiting the participants' arms movements.

7. Conclusion

This paper assessed the effects of selective environmental sound attenuation on operators' task and safety performance, stress, attention, and task engagement. The attenuation filter in the FS condition positively affected operators' performance, stress, attention, and task engagement during teleoperated demolition compared to the US condition. Although the study failed to identify significant differences for some task performance metrics (i.e., demolition performance and SA), it identified positive effects of sound attenuation (FS) on safety metrics (fewer collisions and faster reaction times) compared to US. The FS condition also resulted in lower stress levels, as measured by the LF/HF ratio, and higher levels of attention and task engagement, as measured by the SCR amplitude means, compared to the US condition. These findings suggest that selective sound attenuation can enhance safety and working conditions in construction teleoperation applications. Furthermore, this approach could benefit other applications beyond demolition, provided that the appropriate sound sources are identified for attenuation.

The study's outcomes and limitations open possibilities for future research. First, validating the study's findings with real teleoperated demolition robots and actual operators is the next key step. Additionally, more research is needed to assess if FS improves task and safety performance over NS. Other possibilities include assessing the implications of adjusting the sound levels of the interface to account for personal preferences. Furthermore, understanding how the requirements for sound levels and types change for different applications and task complexities is another direction that begs exploring. Finally, there is a need to investigate the effectiveness of including other sound signals (e.g., sonification for collision warnings) in addition to environmental sound on the operators' performance and psychological and physiological responses.

CRediT authorship contribution statement

Patrick Borges Rodrigues: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Burcin Becerik-Gerber:** Writing - Review & Editing, Supervision. **Lucio Soibelman:** Writing - Review & Editing, Supervision. **Gale M. Lucas:** Methodology, Writing - Review & Editing. **Shawn C. Roll:** Methodology, Writing - Review & Editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Patrick Borges Rodrigues, Burcin Becerik-Gerber, Lucio Soibelman, Gale M. Lucas, Shawn C. Roll reports financial support was provided by National Science Foundation. Burcin Becerik-Gerber, Gale M. Lucas, Shawn C. Roll reports financial support was provided by US Army Research Office. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Science Foundation under Grant No. 2222572 and the U.S. Army Research Office under Grant No. W911NF-20-2-0053. Any opinion, content, or information presented does not necessarily reflect the positions or the policies of the National Science Foundation's or the U.S. Army Research Office's views, and no official endorsement should be inferred.

Data availability

Data will be made available on request.

References

[1] X. Yang, Y. Wang, R. Zhang, Y. Zhang, Physical and psychoacoustic characteristics of typical noise on construction site: "how does noise impact construction workers' experience?", Front. Psychol. 12 (2021) 707868 https://doi.org/10.3389/fpsyg.2021.707868.

[2] K. Elelu, T. Le, C. Le, Collision hazard detection for construction worker safety using audio surveillance, J. Constr. Eng. Manag. 149 (2023) 04022159, https://doi.org/10.1061/JCEMD4.COENG-12561.

[3] J.W. Hinze, J. Teizer, Visibility-related fatalities related to construction equipment, Saf. Sci. 49 (2011) 709–718, https://doi.org/10.1016/j.ssci.2011.01.007.

[4] P. Deshaies, R. Martin, D. Belzile, P. Fortier, C. Laroche, T. Leroux, H. Nélisse, S.-A. Girard, R. Arcand, M. Poulin, M. Picard, Noise as an explanatory factor in work-related fatality reports, Noise Health 17 (2015) 294–299, https://doi.org/10.4103/1463-1741.165050.

[5] M. Mir, F. Nasirzadeh, H. Bereznicki, P. Enticott, S. Lee, A. Mills, Construction noise effects on human health: evidence from physiological measures, Sustain. Cities Soc. 91 (2023) 104470, https://doi.org/10.1016/j.scs.2023.104470.

[6] J. Ke, M. Zhang, X. Luo, J. Chen, Monitoring distraction of construction workers caused by noise using a wearable electroencephalography (EEG) device, Autom. Constr. 125 (2021) 103598, https://doi.org/10.1016/j.autcon.2021.103598.

[7] D.J. Rea, S.H. Seo, Still not solved: a call for renewed focus on user-centered teleoperation interfaces, Front. Robot AI 9 (2022) 704225, https://doi.org/10.3389/frobt.2022.704225.

[8] G. Yamauchi, T. Hashimoto, M. Yamada, S. Yuta, Factors leading to reduced construction productivity in unmanned construction, in: I. Brilakis (Ed.), Proceedings of the 41st International Symposium on Automation and Robotics in Construction, International Association for Automation and Robotics in Construction (IAARC), 2024, pp. 236–242, https://doi.org/10.22260/ISARC2024/0032.

[9] J. Morinaga, Teleoperation System for Ultra-large Hydraulic Excavators, Komatsu Technical Report 68, 2022, pp. 8–15. https://www.komatsu.jp/en/-/media/home/aboutus/innovation/technology/techreport/2022/en/175-e02.pdf?rev=3123acf0d59d4266a066dbc184024217&hash=7465BC2C2126BB257136BD7E1002A41E (accessed April 1, 2024).

[10] J.S. Lee, Y. Ham, H. Park, J. Kim, Challenges, tasks, and opportunities in teleoperation of excavator toward human-in-the-loop construction automation, Autom. Constr. 135 (2022) 104119, https://doi.org/10.1016/j.autcon.2021.104119.

[11] Caterpillar Inc, CAT Command: Remote Control Console and Station. https://s7d2.scene7.com/is/content/Caterpillar/CM20220722-10737-d9a0d, 2022 (accessed April 1, 2024).

[12] International Concrete Repair Institute (ICRI), "Can You Hear Me Now?" How to Preserve Hearing in the Concrete and Masonry Repair Environment, Concrete Repair Bulletin (CRB), 2012, pp. 10–15. https://www.icri.org/wp-content/uploads/2023/12/CRBJanFeb12_SafetySolutions.pdf (accessed May 6, 2024).

[13] H.S. Chan, Occupational noise exposure; criteria for a recommended standard, in: DHHS Publication; No. (NIOSH) 98-126, 1998. https://stacks.cdc.gov/view/cdc/6376 (accessed November 2, 2023).

[14] K. Dang, K. Elelu, T. Le, C. Le, Augmented hearing of auditory safety cues for construction workers: a systematic literature review, Sensors 22 (2022) 9135, https://doi.org/10.3390/s22239135.

[15] P. Bremner, T.J. Mitchell, V. McIntosh, The impact of data sonification in virtual reality robot teleoperation, Front. Virtual Real 3 (2022) 904720, https://doi.org/10.3389/frvir.2022.904720.

[16] E. Triantafyllidis, C. Mcgreavy, J. Gu, Z. Li, Study of multimodal interfaces and the improvements on teleoperation, IEEE Access 8 (2020) 78213–78227, https://doi.org/10.1109/ACCESS.2020.2990080.

[17] A.R. Peon, D. Prattichizzo, Reaction times to constraint violation in haptics: comparing vibration, visual and audio stimuli, in: 2013 World Haptics Conference (WHC), IEEE, 2013, pp. 657–661, https://doi.org/10.1109/WHC.2013.6548486.

[18] Y. Nagai, S. Tsuchiya, T. Iida, S. Kimura, Audio feedback system for teleoperation experiments on engineering test satellite VII system design and assessment using eye mark recorder for capturing task, IEEE Trans. Syst. Man Cybern. Syst. Hum. 32 (2002) 237–247, https://doi.org/10.1109/TSMCA.2002.1021111.

[19] T. Morita, Y. Zhu, T. Aoyama, M. Takeuchi, K. Yamamoto, Y. Hasegawa, Auditory feedback for enhanced sense of agency in shared control, Sensors 22 (2022) 9779, https://doi.org/10.3390/s22249779.

[20] J. Harder, M. Wilde, J. Ventura, M. Dziura, Acoustic telepresence for spacecraft proximity operations, in: 2016 IEEE Aerospace Conference, 2016, pp. 1–12, https://doi.org/10.1109/AERO.2016.7500670.

[21] M. Wang, S.L. Lyckvi, C. Chen, P. Dahlstedt, F. Chen, Using advisory 3D sound cues to improve drivers' performance and situation awareness, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2017, pp. 2814–2825, https://doi.org/10.1145/3025453.3025634.

[22] N. Khenak, J. Vézien, D. Théry, P. Bourdot, Spatial presence in real and remote immersive environments and the effect of multisensory stimulation, Presence Teleop. Virt. 27 (2018) 287–308, https://doi.org/10.1162/pres_a_00332.

[23] R. Secoli, M.-H. Milot, G. Rosati, D.J. Reinkensmeyer, Effect of visual distraction and auditory feedback on patient effort during robot-assisted movement training after stroke, J. Neuroeng. Rehabil. 8 (2011) 21, https://doi.org/10.1186/1743-0003-8-21.

[24] G.A. de Souza, L.A. Torres, V.S. Dani, D.S. Villa, A.T. Larico, A. Maciel, L. Nedel, Evaluation of visual, auditory and vibro-tactile alerts in supervised interfaces, in: 2018 20th Symposium on Virtual and Augmented Reality (SVR), IEEE, 2018, pp. 163–169, https://doi.org/10.1109/SVR.2018.00033.

[25] S. Oviatt, Theoretical foundations of multimodal interfaces and systems, in: S. Oviatt, B. Schuller, P.R. Cohen, D. Sonntag, G. Potamianos, A. Krüger (Eds.), The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1, ACM, 2017, pp. 19–50, https://doi.org/10.1145/3015783.3015786.
[26] K.-C. Siu, I.H. Suh, M. Mukherjee, D. Oleynikov, N. Stergiou, The impact of environmental noise on robot-assisted laparoscopic surgical performance, Surgery 147 (2010) 107–113, https://doi.org/10.1016/j.surg.2009.08.010.
[27] M.J.M. Dunn, B.R.C. Molesworth, T. Koo, G. Lodewijks, Measured effects of workload and auditory feedback on remote pilot task performance, Ergonomics 65 (2022) 886–898, https://doi.org/10.1080/00140139.2021.2003870.
[28] X. Lu, S. Davis, How sounds influence user safety decisions in a virtual construction simulator, Saf. Sci. 86 (2016) 184–194, https://doi.org/10.1016/j.ssci.2016.02.018.
[29] D. Worrall, Sonification: An overview, in: D. Worrall (Ed.), Sonification Design: From Data to Intelligible Soundfields, 1st ed, Springer Cham, Cham, 2019, pp. 23–54, https://doi.org/10.1007/978-3-030-01497-1_2.
[30] X. Lu, S. Davis, Priming effects on safety decisions in a virtual construction simulator, Eng. Constr. Archit. Manag. 25 (2018) 273–294, https://doi.org/10.1108/ECAM-05-2016-0114.
[31] J. Hong, J. Choi, J. Lee, S. Cho, T. Hong, S. Han, H.S. Park, D.-E. Lee, Virtual reality-based analysis of the effect of construction noise exposure on masonry work productivity, Autom. Constr. 150 (2023) 104844, https://doi.org/10.1016/j.autcon.2023.104844.
[32] A. Burova, J. Mäkelä, T. Keskinen, P. Kallioniemi, K. Ronkainen, M. Turunen, The role of audio feedback and gamification elements for remote boom operation, Multimod. Technol. Interact. 8 (2024) 69, https://doi.org/10.3390/mti8080069.
[33] W. Xiong, X. Xu, L. Chen, J. Yang, Sound-based construction activity monitoring with deep learning, Buildings 12 (2022) 1947, https://doi.org/10.3390/buildings12111947.
[34] T.C. Morata, C.L. Themann, R.F. Randolph, B.L. Verbsky, D.C. Byrne, E.R. Reeves, Working in noise with a hearing loss: perceptions from workers, supervisors, and hearing conservation program managers, Ear Hear. 26 (2005) 529–545, https://doi.org/10.1097/01.aud.0000188148.97046.b8.
[35] Z. Mnasri, S. Rovetta, F. Masulli, Anomalous sound event detection: a survey of machine learning based methods and applications, Multimed. Tools Appl. 81 (2022) 5537–5586, https://doi.org/10.1007/s11042-021-11817-9.
[36] B. Sherafat, A. Rashidi, S. Asgari, Sound-based multiple-equipment activity recognition using convolutional neural networks, Autom. Constr. 135 (2022) 104104, https://doi.org/10.1016/j.autcon.2021.104104.
[37] Z. Liu, S. Li, A sound monitoring system for prevention of underground pipeline damage caused by construction, Autom. Constr. 113 (2020) 103125, https://doi.org/10.1016/j.autcon.2020.103125.
[38] A. Mansoor, S. Liu, G.M. Ali, A. Bouferguene, M. Al-Hussein, I. Hassan, Keyword identification framework for speech communication on construction sites, in: Modular and Offsite Construction (MOC) Summit Proceedings, 2022, pp. 106–113, https://doi.org/10.29173/mocs271.
[39] W. Zaharuddin, I. Paraskevas, C. Liu, Accident avoidance importance for building demolition, in: T. Cooke, M. Turner, H. Lingard (Eds.), CIB W099 2009: Working Together: Planning, Designing and Building a Healthy and Safe Construction Industry, RMIT University, Melbourne, Australia, 2009, pp. 9–14. http://www.irbnet.de/daten/iconda/CIB20285.pdf (accessed February 24, 2022).
[40] D.J. Patel, D.A. Patel, Assessing the potential health hazards of workers of demolition sector, in: Construction Research Congress, 2020, pp. 414–423, https://doi.org/10.1061/9780784482872.045.
[41] P.B. Rodrigues, B. Becerik-Gerber, L. Soibelman, G.M. Lucas, S.C. Roll, Virtual environment for studying the effects of operational and environmental sounds on teleoperated demolition, in: Computing in Civil Engineering 2023, American Society of Civil Engineers, Reston, VA, 2024, pp. 54–61, https://doi.org/10.1061/9780784485231.007.
[42] T.M. Marteau, H. Bekker, The development of a six-item short-form of the state scale of the Spielberger state—trait anxiety inventory (STAI), Br. J. Clin. Psychol. 31 (1992) 301–306, https://doi.org/10.1111/j.2044-8260.1992.tb00997.x.
[43] S.G. Hart, L.E. Staveland, Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, in: P.A. Hancock, N. Meshkati (Eds.), Advances in Psychology, North-Holland, 1988, pp. 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.
[44] Centers for Disease Control and Prevention (CDC), What Noises Cause Hearing Loss? Centers for Disease Control and Prevention (CDC), 2022. https://www.cdc.gov/nceh/hearing_loss/what_noises_cause_hearing_loss.html (accessed November 2, 2023).
[45] BAFX Products, BAFX3370 DB Decibel Meter / Sound Level Reader, BAFX Products, 2021. https://bafxpro.com/products/decibel-meter-sound-level-reader (accessed November 2, 2023).
[46] Y.M. Xiao, Z.M. Wang, M.Z. Wang, Y.J. Lan, The appraisal of reliability and validity of subjective workload assessment technique and NASA-task load index, Chin. J. Ind. Hyg. Occup. Dis. 23 (2005) 178–181, https://doi.org/10.3760/cma.j.issn.1001-9391.2005.03.007.
[47] M.R. Endsley, Situation awareness global assessment technique (SAGAT), in: Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, Dayton, OH, USA vol. 3, 1988, pp. 789–795, https://doi.org/10.1109/NAECON.1988.195097.
[48] M.R. Endsley, Direct measurement of situation awareness: Validity and use of SAGAT, in: M.R. Endsley, D.J. Garland (Eds.), Situation Awareness Analysis and Measurement, 1st ed, CRC Press, New York, 2000, pp. 131–157, https://doi.org/10.1201/b12461.
[49] A. Tluczek, J.B. Henriques, R.L. Brown, Support for the reliability and validity of a six-item state anxiety scale derived from the state-trait anxiety inventory, J. Nurs. Meas. 17 (2009) 19–28, https://doi.org/10.1891/1061-3749.17.1.19.
[50] R. Jouvent, C. Bungener, P. Morand, V. Millet, S. Lancrenon, M. Ferreri, Distinction between anxiety state/trait in general practice: a descriptive study, Encephale 25 (1999) 44–49. https://pubmed.ncbi.nlm.nih.gov/10205733/ (accessed November 2, 2023).
[51] H.-G. Kim, E.-J. Cheon, D.-S. Bai, Y.H. Lee, B.-H. Koo, Stress and heart rate variability: a meta-analysis and review of the literature, Psychiatry Investig. 15 (2018) 235–245, https://doi.org/10.30773/pi.2017.08.17.
[52] Polar Electro, Polar H10. https://www.polar.com/us-en/sensors/h10-heart-rate-sensor/, 2022 (accessed November 2, 2023).
[53] M.P. Tarvainen, J.-P. Niskanen, J.A. Lipponen, P.O. Ranta-aho, P.A. Karjalainen, Kubios HRV – heart rate variability analysis software, Comput. Methods Prog. Biomed. 113 (2014) 210–220, https://doi.org/10.1016/j.cmpb.2013.07.024.
[54] R.E. Kleiger, P.K. Stein, J.T. Bigger, Heart rate variability: measurement and clinical utility, Ann. Noninvasive Electrocardiol. 10 (2005) 88–101, https://doi.org/10.1111/j.1542-474X.2005.10101.x.
[55] F. Shaffer, J.P. Ginsberg, An overview of heart rate variability metrics and norms, Front. Public Health 5 (2017) 258, https://doi.org/10.3389/fpubh.2017.00258.
[56] J. Tindle, P. Tadi, Neuroanatomy, Parasympathetic Nervous System, StatPearls Publishing, Treasure Island (FL), 2022. https://www.ncbi.nlm.nih.gov/books/NBK553141/ (accessed January 31, 2024).
[57] A. Pecchinenda, The affective significance of skin conductance activity during a difficult problem-solving task, Cognit. Emot. 10 (1996) 481–504, https://doi.org/10.1080/026999396380123.
[58] G. Turpin, F. Schaefer, W. Boucsein, Effects of stimulus intensity, risetime, and duration on autonomic and behavioral responding: implications for the differentiation of orienting, startle, and defense responses, Psychophysiology 36 (1999) 453–463, https://doi.org/10.1111/1469-8986.3640453.
[59] J.J. Braithwaite, D.G. Watson, R. Jones, M. Rowe, A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments (Revised Version: 2.0), Birmingham, England. https://www.birmingham.ac.uk/documents/college-les/psych/saal/guide-electrodermal-activity.pdf, 2015.
[60] Empatica Inc, E4 Sensors. https://www.empatica.com/research/e4/, 2018 accessed November 2, 2023).
[61] M. Nabian, Y. Yin, J. Wormwood, K.S. Quigley, L.F. Barrett, S. Ostadabbas, An open-source feature extraction tool for the analysis of peripheral physiological data, IEEE J. Transl. Eng. Health Med. 6 (2018) 1–11, https://doi.org/10.1109/JTEHM.2018.2878000.
[62] G. Cosoli, A. Poli, L. Scalise, S. Spinsante, Measurement of multimodal physiological signals for stimulation detection by wearable devices, Measurement 184 (2021) 109966, https://doi.org/10.1016/j.measurement.2021.109966.
[63] M.J. Blanca, J. Arnau, F.J. García-Castro, R. Alarcón, R. Bono, Non-normal data in repeated measures ANOVA: impact on type I error and power, Psicothema 35 (2023) 21–29, https://doi.org/10.7334/psicothema2022.292.
[64] U. Knief, W. Forstmeier, Violating the normality assumption may be the lesser of two evils, Behav. Res. Methods 53 (2021) 2576–2590, https://doi.org/10.3758/s13428-021-01587-5.
[65] S. Lee, D.K. Lee, What is the proper way to apply the multiple comparison test? Korean J. Anesthesiol. 71 (2018) 353–360, https://doi.org/10.4097/kja.d.18.00242.
[66] L.J. Julian, Measures of anxiety: state-trait anxiety inventory (STAI), Beck anxiety inventory (BAI), and hospital anxiety and depression scale-anxiety (HADS-A), Arthritis Care Res. 63 (2011) S467–S472, https://doi.org/10.1002/acr.20561.
[67] M. Iwanaga, A. Kobayashi, C. Kawasaki, Heart rate variability with repetitive exposure to music, Biol. Psychol. 70 (2005) 61–66, https://doi.org/10.1016/j.biopsycho.2004.11.015.
[68] S. Hammoud, R. Karam, R. Mourad, I. Saad, M. Kurdi, Stress and heart rate variability during university final examination among Lebanese students, Behav. Sci. 9 (2018) 3, https://doi.org/10.3390/bs9010003.
[69] T. Zetterman, R. Markkula, T. Miettinen, E. Kalso, Heart rate variability responses to cognitive stress in fibromyalgia are characterised by inadequate autonomous system stress responses: a clinical trial, Sci. Rep. 13 (2023) 700, https://doi.org/10.1038/s41598-023-27581-9.
[70] M. Shibagaki, T. Yamanaka, Attention of preschool children: temporal variation of electrodermal activity during auditory stimulation, Percept. Mot. Skills 72 (1991) 1115–1124, https://doi.org/10.2466/pms.1991.72.3c.1115.
[71] P.A. Hancock, J.S. Warm, A dynamic model of stress and sustained attention, Human Factors 31 (1989) 519–537, https://doi.org/10.1177/001872088903100503.
[72] R.J. Tafalla, G.W. Evans, Noise, physiology, and human performance: the potential role of effort, J. Occup. Health Psychol. 2 (1997) 148–155, https://doi.org/10.1037/1076-8998.2.2.148.
[73] Y. Sasaki, S. Masunaga, S. Thompson, S. Kagami, H. Mizoguchi, Sound localization and separation for mobile robot tele-operation by tri-concentric microphone array, J. Rob. Mechatron. 19 (2007) 281–289, https://doi.org/10.20965/jrm.2007.p0281.
[74] P. Adami, P.B. Rodrigues, P.J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk, G. Lucas, Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation, Adv. Eng. Inform. 50 (2021) 101431, https://doi.org/10.1016/j.aei.2021.101431.