

Can VR training complement existing initial training and professional development for hoist operators?

Authors: Claire M. Oakley, Marie Juanchich, Motonori Yamaguchi, Ian Bonthrone, David Bazlinton, Trevor Abraham, Miroslav Sirota

Date: 21.09.2026

Version: FINAL REPORT

This work was supported by the University of Essex Innovation Fund (RCP20079).

Authors and affiliations

Claire M. Oakley
Department of Psychology, University of Essex, Colchester, United Kingdom.

Marie Juanchich
Department of Psychology, University of Essex, Colchester, United Kingdom.

Motonori Yamaguchi
Interdisciplinary Faculty of Science and Engineering, Shimane University, Matsue, Japan.

Ian Bonthrone
Air Rescue Group, Harwich, United Kingdom.

David Bazlinton
Air Rescue Group, Harwich, United Kingdom.

Trevor Abraham
Air Rescue Group, Harwich, United Kingdom.

Miroslav Sirota
Department of Psychology, University of Essex, Colchester, United Kingdom.

Contents

List of Abbreviations 4

Brief Definitions 4

Executive Summary 5

Key Findings 5

Public-facing Summary 8

Chapter 1. Scoping Review 10

Background 10

Literature Search Approach 12

What do we mean by Virtual Reality, Immersion, and Fidelity. 14

What is the value of VR training in hoist operations? 16

What technical skills are relevant to hoist operations that VR training should develop? 16

What non-technical skills are relevant to hoist operations that VR training should develop? 18

Is VR simulation effective to train non-technical skills for hoisting? 21

Can VR Training Simulation enhance decision-making? 22

Can hoist training be optimised to account for trainees' characteristics? 29

How could we standardise and optimise VR hoist operator training? 24

Conclusion 30

Chapter 2. VR Training Evaluation – Testing the Benefits and Limitations for Hoist Operations **31**

Survey Development 31

Measures 31

Demographics 33

Preliminary Results 34

Conclusion 44

Chapter 3. Recommendations for Phase 2+ Research **46**

1. Validate Fidelity, Construct Validity, and Real-World Skill Transfer 46

2. Implement Reliable, Objective Measures of Cognitive Load and Stress 47

3. Develop Adaptive, Experience-Tailored Training Pathways 47

4. Profile Individual Psychological Predictors of VR Learning 47

5. Validate Specific Non-Technical Skills Assessment Frameworks 48

References **49**

List of Abbreviations

CAVE	Cave automatic virtual environments.
CRM	Crew Resource Management.
EASA	European Union Aviation Safety Agency.
HEMS	Helicopter Emergency Medical Services
MR	Mixed Reality.
SAR	Search and Rescue.
VR	Virtual Reality

Brief Definitions

Virtual Reality

Virtual reality is a three-dimensional computer-generated environment that replicates the physical world, allowing users to navigate and interact with in real time.

Mixed Reality

Mixed reality is an immersive (e.g., using head-mounted displays) extension of a VR environment in which physical and virtual objects co-exist. Users can see and physically interact with physical and virtual objects simultaneously, receiving tactile feedback that replicates the real-world experience.

Augmented Reality

Augmented reality overlays computer-generated images, text, or visual data onto a real-world viewing window in real-time.

CAVE

CAVE is a room-sized immersive alternative to immersive VR in which three-dimensional images are projected onto walls, floors, and ceilings within an enclosed space.

Executive Summary

This report, funded through the University of Essex's Innovation Fund, presents a Phase 1 scoping and evidence review evaluating the efficacy of virtual and mixed-reality (VR/MR) simulation for helicopter hoist operator training. The project sought to evaluate the current VR-related hoist operator and broader helicopter emergency services training literature in respect of four key goals outlined in the original application: demonstrating training value, identifying psychological mechanisms of impact, exploring predictors of performance, and optimising training design. While the project planned to analyse raw objective simulator and cognitive testing data, this was not feasible due to data formatting constraints and access issues. Instead, the project team supported Air Rescue UK in designing and implementing a new intra-day trainee evaluation survey. Air Rescue incorporated this intra-day survey into their standard training days while the project was ongoing. This report synthesises a comprehensive scoping review of the literature alongside the quantitative and qualitative analysis of the intra-day survey.

Key Findings

Demonstrating Training Value.

The literature establishes VR and MR as safe, cost-effective, and highly scalable alternatives for practising complex, high-risk, or rare scenarios in helicopter-based and other emergency operations. VR/MR-based training is particularly valuable for training scenarios that cannot be experienced in live training, for instance, severe weather, night flying, and emergency system failures, without endangering the crew or aircraft. Trainee feedback collected through Air Rescue UK's training evaluation survey corroborated this. Overwhelmingly, hoist operators valued the MR simulations as an acceptable complement to the standard curriculum that could reduce live-flight requirements. Qualitatively, hoist operators highly praised the ability to safely execute critical emergency procedures, such as winch runaways, cable fouls, and simulated hoist cuts. However, a major gap exists in the literature regarding empirical evidence confirming that technical and non-technical skills acquired in VR or MR-based hoist operator training are successfully transfer to real-world live hoisting or HEMS operations.

Psychological Mechanisms of VR Training Effectiveness

Research from the broader literature shows that high-fidelity virtual and mixed reality simulations train critical non-technical skills through two primary psychological mechanisms: affective fidelity (inducing realistic psychological stress to promote stress inoculation) and cognitive load calibration (matching task complexity to trainee processing capacity to maximise engagement). By optimising affective fidelity and cognitive load, trainees can be trained to enhance decision-making, communication, and situational awareness under pressure. Survey results indicated that the MR hoisting training produced manageable subjective workloads for almost all training missions. Still, there is room for improvement for the most experienced hoist operators, who often use MR-based training for practising emergency scenarios and may benefit from higher subjective workloads. Overall, the survey indicated that there was a strong correlation between the subjective workload a trainee experienced and their perceived educational benefit, and the importance of individual workload factors (e.g., their perceived performance in the simulation, time pressure, effort) differed by hoisting experience level. Together these indicate that cognitive load and perceived challenge were key drivers of perceived training value.

Optimising Training Design

Responses to the Air Rescue training survey indicated that training outcomes differed by experience level, therefore MR-based hoist training may need to account for trainee experience level to optimise learning for future trainees. Evidence from the literature review suggests that novice operators may benefit from lower cognitive-load introductory scenarios and structured pauses in which trainees verbalise their evaluation of the scenario and hoisting or approach plan. Similarly, scaffolded gaze training can support trainees in efficiently identifying and attending to important environmental objects such as snag points. This approach allows trainees to build foundational schemas before progressively introducing complex, high-workload emergencies. In contrast, experienced hoist operators require sufficiently demanding, time-pressured scenarios to match their higher capacity. Real-time calibration of this difficulty is best supported by objective, non-invasive markers of cognitive workload. Gaze fixation patterns and vocal stress indicators could be evaluated as cognitive workload markers as wearable physiological sensors are prone to confounds, such as the physical exertion involved in cable handling and clothing worn during hoisting sessions. However, two design questions remain unanswered by the current literature: how hoist training should be scheduled (e.g., spaced versus massed practice, or the optimal balance of MR and live-flight hours),

and which individual characteristics beyond hoisting experience, such as personality, spatial ability, or prior VR/MR exposure, should inform how training is tailored to the trainee.

Predictors of Learning and Performance

Responses to the Air Rescue UK training survey indicated that a trainee's prior real-world hoisting experience predicted their perceived learning experience, meaning training approaches may need to be optimised to match the trainees' needs. For example, highly experienced operators (15+ years) reported the greatest learning benefits from shorter, highly mentally challenging, and time-pressured emergency scenarios, whereas there were low or non-significant correlations between subjective workload and learning outcomes for a group of mixed-experience hoist operators. This finding suggests that the predictors of learning outcomes may differ among learners with differing experience levels; further data collection is required to clarify these associations.

Concluding Remarks

Together, the findings from the literature review and the survey point to five priority areas for Phase 2 research and beyond (see *Recommendations for Phase 2+ Research*: validating simulator fidelity and real-world skill transfer; implementing reliable, objective measures of cognitive workload and stress; developing adaptive, experience-tailored training pathways and profiling the individual psychological predictors that inform the training pathways; and validating hoisting-specific non-technical skills assessment frameworks. Optimal training delivery scheduling is also flagged as a valuable question for a future phase, though this may require a partner organisation to be able to compare live-only against combined MR/live training pathways.

Public-facing Summary

Beyond Live Flight: The Scientific Evidence Behind Air Rescue UK's Next-Generation Hoist Simulator, in Partnership with the University of Essex

Air Rescue UK, a specialist provider supporting helicopter hoist operators in Europe, Middle East, and Africa, has partnered with researchers from the Department of Psychology, University of Essex to scientifically evaluate and optimise their cutting-edge mixed-reality simulator for helicopter hoist operator training. Supported by the University of Essex Innovation Fund, this multi-phase collaborative project aims to ensure that Air Rescue UK's innovative training pathways are grounded in rigorous academic research, specifically focusing on the psychological characteristics and cognitive demands that underpin expert hoisting performance under extreme pressure.

Helicopter hoist operations are highly complex and dangerous, requiring operators to manage aircraft movement, environmental hazards, crew communication, and technical rescue tasks simultaneously. Currently, much of the industry relies heavily on costly live-flight training, which exposes crews and aircraft to the high risks they are attempting to manage. Live-flight training incurs significant financial expenses - consuming fuel, adding wear-and-tear to the helicopter, and fundamentally restricting a trainee's ability to practice emergencies safely. Because instructors cannot intentionally place learners in highly dangerous situations (e.g., rough seas, severe winds, or critical equipment failures), operators relying solely on live training miss the opportunity to develop the well-trained, swiftly judged, and smoothly executed life-saving responses needed for the most extreme rescue scenarios.

To solve these critical industry challenges, Air Rescue UK offers an advanced mixed-reality simulator that provides an engaging, cost-effective, and safe alternative to traditional live flights. The simulator allows operators to repeatedly practice high-risk, time-critical emergency procedures, for example winch-runaways, cable-fouls, engine failures during hoisting, and simulated hoist cuts, that are too dangerous to rehearse in a live aircraft. The system provides physical realism by incorporating a physical hoist cable with accurate tension and directionality, allowing operators to physically feel and manage cable plumb, rotor wash, and shifting loads just as they would in real-life missions. By utilising this immersive technology, operators can execute back-to-back training missions in challenging environmental conditions without the logistical constraints of refuelling or burning expensive live flight hours.

In this initial phase, the training evaluation survey responses collected by Air Rescue UK demonstrate the value and high trainee acceptance of their MR hoist operator training system. Feedback from hoist operators is overwhelmingly positive, with trainees confirming that the simulator is a highly useful complement to the standard curriculum that can effectively reduce the amount of high-risk live training required. Up to 92% of post-training evaluations indicated trainees had learned something new or refined their skills. Furthermore, trainees reported excellent visual quality and rated the mental demands of their training missions as well-matched to their skill level overall.

Through this ongoing academic partnership, Air Rescue UK and the University of Essex are using these early findings to guide Phase 2 research into training pathways to tailor to individual hoist operators' experience levels. For operators currently relying solely on live-flight hours, integrating Air Rescue UK’s evidence-based simulation provides a highly cost-effective way to safely increase training volume, elevate crew resilience, and ensure operators are prepared for the unpredictable realities of modern aerial rescue.

Chapter 1. Scoping Review

Background

Helicopter hoist operations retrieve patients or personnel from locations that are inaccessible by ground transport or where challenging terrain prevents the helicopter from landing (Pietsch et al., 2021). For instance, crews may be required to extract mountain climbers from steep, vertical alpine environments, or retrieve victims from a sinking vessel where there are numerous snag points for the hoist cable (King et al., 2022). These operations are complex, requiring the crew to simultaneously manage the aircraft and technical rescue and balance rescue responsibilities against aircraft safety (Hamlet et al., 2020; Pietsch et al., 2016). While the pilot is in command of the helicopter, the hoist operator leads the rescue mission in many respects, acting as the eyes for the pilot who cannot see directly underneath the helicopter. The hoist operator directs the pilot, requesting changes in direction and speed to position the helicopter above the rescue target. To ensure safety and effective operation, the hoist operator must account for rotor wash (airflow generated by the helicopter's main rotor), environmental hazards, weather conditions, and continually monitor for obstacles during each hoist operation (King et al., 2022).

Despite the number of skills required and their critical importance, hoisting operations – and the training of hoist operators - vary considerably across jurisdictions and operational models. In military and dedicated operations, such as helicopter operations serving offshore platforms or installations, hoisting is typically performed by professional, highly specialised hoist operators whose primary role and expertise revolve around managing the winch system and cable dynamics (King et al., 2022). Conversely, in many civilian Helicopter Emergency Medical Services (HEMS), the hoist operator role is fulfilled by medical personnel, such as paramedics (Hamlet et al., 2020; Pietsch et al., 2016), who receive supplementary technical training to act as the hoist operator while a physician is lowered to the scene to treat the patient (Pietsch et al., 2020). Although the high levels of technical and cognitive skills are required for helicopter hoisting operations, there are no industry-wide,

standardised training guidelines outlining the skills or skill levels that are deemed safe and proficient within the broad range of operating environments a hoist operator can work in (King et al., 2022; Pietsch et al., 2020).

Hoist training most often follows either an apprenticeship model or a trial-and-error approach during live flight operations (King et al., 2022; Pietsch et al., 2016). Furthermore, regulatory bodies, such as the European Union Aviation Safety Agency (EASA), require hoist operators to undertake initial training, with regular additional training to maintain currency (EASA, 2024), where currency refers to mandated role-specific experience, such as a minimum number of hoist operations over-water, or at high-altitude, that must have taken place within set time period. Fulfilling these requirements through live helicopter flights is expensive as each sortie adds flying hours to life-limited helicopter parts after which they must be changed or overhauled, uses fuel, and produces CO2. It also exposes crews to the risks they are being trained to manage, which could result in fatalities or loss of the aircraft during live training depending on the decision the hoist operator makes (King et al., 2022; Pietsch et al., 2020). Because of the need to avoid these safety risks, real-life training sorties are insufficient to develop well-rounded hoist operators, preventing learners from developing helpful schemas and effective responses to dangerous situations – such as high winds, cable swing, difficult-to-access landscapes, rough sea conditions – even if they will have to face those risks in the future.

To overcome the lack of standardisation, while mitigating the risks of high-cost live-flight practice and providing experience of high-risk situations, organisations such as Air Rescue UK are increasingly using MR simulators in hoist operator training pathways. Simulation-based training has been common practice in pilot training for many years. Unlike head-mounted VR or MR systems that immerse the user in a 360-degree virtual space, traditional flight simulators rely on fixed physical cockpits and screen projections tailored to pilot instrumentation. Scientific evidence shows that flight simulator training, in combination with real-life aircraft training, is more effective than real-life aircraft training alone (Hays et al., 1992). Nevertheless, full cockpit flight simulators are expensive to purchase and, as digital replicas of a flight deck, do not offer an immersive solution. Flight simulation training, therefore, often focuses on procedural matters, reducing in-aircraft hours as a result. However, the lack of visual immersion means they remain unsuitable for many tasks – VR-related technologies are therefore being investigated as a cost-effective addition and, potentially, eventual replacement for traditional flight simulators (Somerville et al., 2025). Can VR also be beneficial for hoist training too?

Given the lack of standardised training, and addressing the broader call for safer, objective training in HEMS operations (Abrahamsen et al., 2015; Hamlet et al., 2023; King et al., 2022; Pietsch et al., 2016), this scoping review assesses the evidence on helicopter-related or simulation-based training to inform future collaborations with Air Rescue UK, with the aim of optimising training design and demonstrating the value of VR-based hoist training. This report synthesises current literature on the use of VR-related technologies for training hoisting and broader HEMS skills across air ambulance, SAR, and emergency response operations. To evaluate existing evidence and guide future implementation, this scoping review addresses four core research questions:

Operational Value & Effectiveness. What is the value and effectiveness of VR simulation training for HEMS and hoist operators compared with existing live or classroom-based strategies, and does it serve as a safe, cost-effective alternative or complement to live training?

Cognitive & Non-Technical Skills. What are the main psychological processes, cognitive mechanisms, and non-technical skills that are actively engaged and developed during effective virtual hoisting operations?

Pathway Optimisation & Standardisation. How can simulation features and design factors be systematically utilised, evaluated, and standardised to build effective, evidence-based training pathways for trainees with different experience levels?

Individual Differences & Tailored Training. Which psychological predictors, learner characteristics, and baseline experience levels influence VR learning outcomes for hoist operators trained using HEMS/SAR simulations, and how can these be used to dynamically tailor training to individual operator needs?

Literature Search Approach

This scoping review was informed through a systematic literature search across four databases—Scopus, Web of Science, PubMed/MEDLINE, and Google Scholar—to identify relevant peer-reviewed and grey literature. The search strategy used a combination of primary keywords related to virtual reality environments (e.g., "virtual reality," "mixed reality," "immersive") and learning outcomes (e.g., "train*," "skill*"), alongside secondary keywords focusing on helicopter emergency operations (e.g., "HEMS," "search and rescue") and specific hoisting terminology (e.g., "winch*,"

"human external cargo"). The initial search yielded 482 references, which were reduced to 398 after removing 84 duplicates. An initial screening of article titles and abstracts discounted 357 articles as not relevant; many of these focused on VR training for medical procedures, proposed specific technologies for VR applications, or were related to VR training for mass casualty events, e.g., mining or earthquake disaster responses, that were deemed less relevant in the context of hoisting training missions. Following this initial review, 40 publications were selected for a comprehensive full-text review and potential information extraction to inform the scoping review. The literature for full-text review was categorised into VR hoisting operations, general HEMS-related literature, and studies informing on psychological processes and VR training effectiveness. From this pool of 41, 25 articles were excluded from the data extraction process due to providing only background information, focusing on VR domains or applications with no obvious synergy with hoisting or HEMS operations, or testing measures that were not related to our knowledge of hoisting operations. This exclusion process resulted in a final cohort of 16 relevant articles that were selected for thorough information extraction. Of these, only 3 articles were both VR and either hoisting or HEMS training related. The remainder of articles reviewed here encompass non-VR HEMS physical simulation-based training (2 articles) that are informative for measuring broader skills used by the helicopter hoisting operations, or from the broader VR emergency-scenarios training literature (e.g., firefighting) that are potentially informative for future training design and evaluation (11 articles).

The report is also supplemented by findings from broader systematic literature reviews and meta-analyses summarising the evidence on the effectiveness of VR-based training on performance, stress, safety-related non-technical skills under high-risk conditions. Some citation chaining was used (both forward and backward) by reviewing the Google Scholar cited by link results and article references lists, to identify any further hoisting or HEMS literature that may be informative for this review. Articles identified through citation chaining focused on identifying non-technical skills measurement for HEMS crew members or the evaluation of the benefits of simulated vs live training; 6 additional articles were included in the review through this method. Additional manual searching was used to identify further evidence on whether VR environments enable riskier decision-making during training given the lack of real-world consequences in simulations and individual differences in VR learning and acceptance as the evidence identified in the review could not fully address these topics.

What do we mean by Virtual Reality, Immersion, and Fidelity.

Virtual Reality Environments and Immersion

Virtual reality is a three-dimensional computer-generated environment that users can navigate and interact with. Fully immersive environments, such as a wearable head-mounted display, allow users to perceive visual and auditory information and perform natural actions when interacting with the virtual environment, like turning their head to look around. **Mixed reality** is an extension of a virtual immersive environment, in which physical objects are available and can be interacted within both the physical and virtual world (Kaplan et al., 2021). For example, mixed reality allows users to virtually see, and physically touch, objects and, in turn, receive tactile feedback. Mixed reality is especially relevant to hoist training as hoist operators rely not only on visual and auditory feedback, but also on haptic (tactile) feedback that provides information about the status of the hoisting operation, such as the external forces acting upon the cable. How hoist operators manage the cable is an important aspect of successful hoist operations (King et al., 2022; King et al., 2020).

Immersion refers to the extent to which the technology surrounds the user with sensory information and allows natural interaction, whereas presence is the user's subjective sense of being in the virtual environment. An environment is described as fully immersive if it enables users to perceive and interact with it as they would in real life (Stefan et al., 2023; Wheeler et al., 2021). In a low-immersion environment, such as a desktop-based VR system with a games controller or keyboard-based controls, users will perform unnatural actions to interact with the virtual environment. An immersive alternative to head-mounted displays are **cave automatic virtual environments** (CAVE) in which three-dimensional images are projected onto walls, floors, and ceilings within an enclosed space. CAVEs provide high-quality visuals and a high-level of immersion but are costly, complex to set up, and can restrict user movement (Scorgie et al., 2024). A final alternative is **augmented reality** in which users see the real-world environment with virtual images and information overlayed in real-time (Kaplan et al., 2021).

Simulation Fidelity and Validity

Simulation fidelity refers to how well the simulation accurately recreates the appearance of the real-world system and elicits user affective states, cognitions, and behaviours that mirror real-world experiences (Evans et al., 2023; Harris et al.,

2020). A simulation must achieve sufficient cognitive, emotional, and perceptual fidelity to ensure that the skills learned in the simulation can transfer successfully to the real world (Harris et al., 2020). While fidelity is often subjectively assessed by users based on how realistic they perceive the simulation to be, and how stressed they were, or how difficult they found the scenario, subjective perceptions and even trainer observations may be inaccurate (Evans et al., 2023). To optimise training outcomes, simulations should sufficiently replicate the appearance (**face validity**), specific perceptual-cognitive demands (**psychological fidelity**), the motor movements and feedback (**ergonomic fidelity**), and the stress-inducing conditions (**affective fidelity**) of the actual environment, while being able to distinguish the performance of experts from novice users (**construct validity**), and to track novice users' improvements (Harris et al., 2020). However, these fidelity types are not equally consequential for training outcomes. Harris et al. (2020) argue that physical fidelity and face validity primarily increase initial user engagement rather than driving skill acquisition, noting that a simulation with low face validity can still be highly effective. Greater physical realism can deepen a user's sense of presence but whether this translates into improved learning or transfer beyond sustaining engagement remains unestablished. Psychological and affective fidelity appear more consequential: perceptual-cognitive skills trained under matched cognitive demands transfer to related real-world tasks regardless of physical realism, and familiarisation with realistic affective states such as stress and anxiety improves subsequent real-world performance under similar conditions (Harris et al., 2020). Consistent with this, Cohen-Hatton & Honey's (2015) three-environment comparison - an immersive VR simulation, a physical fire-ground exercise, and a live burn - found highly consistent decision-making patterns and training benefit across all three, despite the substantial differences in physical realism between them. Objectively measuring affective fidelity by assessing autonomic nervous system arousal, using measures such as heart rate, skin conductance (physiologic fidelity: Evans et al., 2023), independent of perceived emotions, allows simulations to be refined to more closely approximate reality and to manage cognitive load, which may also improve learning outcomes (e.g., Valente et al., 2025).

Higher simulator fidelity is assumed to improve transfer of learning, better preparing learners for the complexities of real-world scenarios, leading to improved operational outcomes (Evans et al., 2023; Harris et al., 2020). The system's technical immersion capabilities provide the sensory input required to provide a strong sense of perceived presence (Harris et al., 2020; Stefan et al., 2023). This perception of presence within the simulated scenario is prerequisite for eliciting realistic affective states, but does not itself ensure fidelity. By achieving high fidelity, meaning that the simulation

accurately simulates a real-world task, the simulator should elicit realistic human behaviours, enabling effective training transfer to the real-world. However, the evidence for transfer of training gains to real-world situations remains limited. Relatively few studies incorporate a real-world transfer assessment into their evaluation (Abich et al., 2021; Gordon et al., 2012; Harris et al., 2020).

What is the value of VR training in hoist operations?

Virtual and mixed reality simulations offer a highly engaging, cost-effective, and safe alternative to live training across a broad range of application domains including medical training, disaster response, military, and offshore operations (Stefan et al., 2023). To demonstrate the value of mixed training for helicopter hoist operations, it is necessary to distinguish between the acquisition and objective measurement of technical and non-technical skills, and to establish construct validity – whether the simulation can accurately distinguish real-world experts from trainees and track performance improvements over time.

What technical skills are relevant to hoist operations that VR training should develop?

Technical hoisting skills refer to the specific psychomotor, procedural, and operational skills required to physically execute a rescue. Two studies have evaluated technical skills captured during mixed-reality simulated hoisting training missions that can be used to discriminate between experienced and novice hoist operators (King et al., 2022; King et al., 2020). In each study, experienced and novice hoist operators (16 and 23 participants respectively) completed hoisting training missions across two different wind conditions, and in second study, over land and sea terrain. Hoist operators interacted with a hoist cable represented in both the physical and virtual environments, as well as a mechanism to raise and lower it. The mixed-reality simulator used in these studies provides haptic feedback by incorporating a physical hoist cable with realistic tension and directionality that coordinates with the virtual display (King et al., 2022; King et al., 2020). This tactile information allows operators to physically feel and evaluate the hoisting operation, providing realistic feedback based on wind conditions, rotor wash, and physical load on the cable, creating a psychomotor feedback loop for managing cable plumb - the displacement of the load at the end of the line. Measures evaluated for distinguishing

between expert and novice users included training mission completion time and the number of helicopter trajectory adjustments made during the approach (conning), as directed by the hoist operator. Other hoist-specific measures tested included cable management metrics including cable plumb, mean cable tension, hand position on the cable, and physical cable displacement by the hoist operator. Both studies reported that experienced hoist operators completed hoisting training missions faster (although this difference was not statistically significant in the first study), managed the hoisting cable more efficiently, and, in the second study, executed fewer conning adjustments (King et al., 2022; King et al., 2020).

The two studies highlight that there are limited differences between novice users and experienced hoist operators in no-wind or low-wind conditions. The differences between these groups became more apparent in windy (25-knot) conditions or in training missions over challenging terrains. For example, King et al. (2022) found that novice users took longer to hoist over land than over the sea, exhibited more cable displacement towards the rear of the aircraft, under higher winds moving forward to aft, and their cable displacement was more variable than experienced hoist operators. Evaluating whether these objective technical skills - such as cable plumb, cable tension, conning adjustments, and hand position - differ between experienced operators and novices serves to establish the construct validity of the simulator (Harris et al., 2020; King et al., 2020, 2022). Demonstrating that real-world expertise predicts superior simulator performance under difficult environmental conditions provides initial evidence that the system captures fundamental psychomotor demands. However, while these studies establish initial construct validity, important measurement gaps remain. King et al. (2020, 2022) did not evaluate the ergonomic or biomechanical fidelity of the system; that is, whether experienced operators subjectively perceive the haptic cable forces and physical movement patterns as matching live-flight hoisting. Furthermore, because control groups consisted of general-population volunteers rather than actual hoist trainees, further work is required to confirm construct validity across realistic trainee cohorts

Identifying measures that successfully distinguish between expert hoist operators and novices can potentially demonstrate that the simulation has good construct validity (Harris et al., 2020). However, while these papers represent initial attempts at establishing construct validity for hoisting training mission simulation, the results are not conclusive. Low sample size may contribute, but while the experienced group were military hoist operators, most of whom were on active duty, the comparison group were general-population volunteers and not drawn from a pool of trainee hoist operators. It is possible that the baseline skills of a group of self-initiating hoisting

trainees may differ from a random sample of volunteers drawn from the general population who do not intend to train as hoist operators. For example, from to Air Rescue UK experience, trainee hoist operators often have flying or military experience or are highly trained medical personnel. Further work is therefore needed to clarify exactly which measures and in which conditions the simulator can consistently distinguish between experienced and novice hoist operators - especially as the literature search did not identify additional evidence on the measurement of technical hoisting skills in VR beyond these two studies. In addition, no studies were found that measured user perceptions of the simulations to evaluate face and physical fidelity, nor did they measure affective (physiologic) fidelity to demonstrate that the training missions tested elicited realistic levels of stress in the participants (Harris et al., 2020). Higher simulation fidelity – specifically the emotional, cognitive and perceptual demands elicited by the simulation - have been linked to faster skills acquisition and training outcomes (Harris et al., 2020). Neither did the review identify any evidence of the transfer of learning to live missions within the hoisting or HEMS contexts.

What non-technical skills are relevant to hoist operations that VR training should develop?

Non-technical skills encompass the cognitive and social competences that complement a hoist operators' technical skills to needed to safely operate the hoist. Despite their importance to ensure safe and effective task performance under pressure, there is no universal agreement on the specific non-technical skills required for hoist operators. Instead, core non-technical skills commonly identified within the broader HEMS literature include situation awareness, decision making, communication, teamwork, leadership, managing stress, and coping with fatigue (Abrahamsen et al., 2015; Pietsch et al., 2016). These core skills are embedded within crew resource management (CRM) training, a multidisciplinary framework originally conceptualised and developed within the aviation industry to reduce human error and increase overall crew effectiveness (Abrahamsen et al., 2015; Hamlet et al., 2020; O'Connor et al., 2008). CRM training is a mandatory component of hoist operator training under the EASA guidelines (EASA, 2024).

Like other emergency and rescue scenarios, and indeed military operations, hoisting training missions are highly dynamic and reactive. This means that hoist operators and their operational crew must continuously assess the situation, which may change as the training mission progresses, adapt their plans and communicate accordingly. Of particular concern is overcommitment to the rescue, resulting in

operational team casualties (e.g., in firefighting: Ash & Smallman, 2008), or loss of or damage to the helicopter. For example, in 1999, two firefighters got lost in a search for two civilians suspected to be in a burning building. Despite the rapidly deteriorating situation, further crews were dispatched into the building to continue the search. The consequences of persisting despite escalating risks resulted in unnecessary additional loss of life (Braddee et al., 2000). Examples like these demonstrate that reactive measures taken during operations must weigh the level of risk against the potential benefit. In firefighting scenarios, responders will take risks to save lives but must also react by ceasing rescue attempts if the risk of harm to themselves and their crew becomes too high (Tissington & Flin, 2005). Similar risk-versus-benefit decisions must be made during HEMS missions, including hoisting training missions (Hamlet et al., 2020; Pietsch et al., 2021). Here, the hoist operator must weigh the risk to the helicopter, crew, and the person being rescued, particularly when unexpected events occur, or, for example, when hoisting over more complex terrain where obstacles may impede access or may be a risk for cable entanglement. An example of this risk occurred during a 2011 US Air Force training mission, where a flight engineer's focussed attention inside the cabin on a simulated hoist malfunction during a life training exercise caused the aircraft to drift undetected, resulting in the rescue device snagging on a ground obstacle. This entanglement tethered the helicopter, leading to a crash that destroyed the aircraft (Plamp, 2011). Non-technical skills training for hoist operators therefore must train for and assess competency across scenarios that require adaptation and reactive decision-making–scenarios that are high risk and unsuitable for experience-based training in live operations (Abrahamsen et al., 2015). Simulated VR-based training offers clear, demonstrable benefits for high-risk scenario training in a range of domains including aviation, firefighting, mining, and the military (Scorgie et al., 2024; Stefan et al., 2023; Wheeler et al., 2021), but it is yet to be tested in the hoisting context. In the hoisting context, it is feasible to design scenarios that test hoist operators' non-technical skills without risking the crew or the helicopter, however the effectiveness of high-risk scenario training for hoist operators is not established.

While the standard CRM skills definition provides a set of foundational safety-related skills, as they were conceptualised for commercial aviation, they do not reflect the skills needed to adjust for the dynamic and reactive nature of helicopter hoist operations. Beyond the CRM skills, hoist operators may also need to be proficient in cognitive readiness, a skill category defined for operations in reactive environments. Cognitive readiness reflects the skills needed to respond rapidly and effectively to unpredictable situations (Hamlet et al., 2020; O'Neil et al., 2014). This skill category encompasses the ability to adapt to change, including the failure of prior plans, the

mental and physical preparedness to respond to unexpected events, psychological resilience to maintain optimal performance under severe stress, and complex problem-solving skills enabling the evaluation of tasks from multiple perspectives to cope with challenging rescue conditions (Hamlet et al., 2020). Through a combination of discussion groups with pilots and technical crew operating in offshore helicopter operations and SAR missions, Hamlet et al. (2023) proposed the HeliNOTS (SAR) and HeliNOTS (O) systems of behavioural markers for assessing non-technical skills in SAR and offshore helicopter operations, respectively. The systems propose a set of positive and negative behaviour markers and a rating scale based on the NOTECHS system for assessing pilots' non-technical skills (e.g., Flin, 2010). Both proposed HeliNOTS systems incorporate communication, leadership / teamwork, situational awareness, and decision-making. They differ by incorporating cognitive readiness into the SAR version and workload management – coping with task demands and prioritisation – in the offshore version (Hamlet et al., 2023). Air Rescue training would necessarily need to address both workload management and cognitive readiness, given that the training Air Rescue offers prepares hoist operators for both HEMS and offshore hoist operations. However, to date, the HeliNOTS systems have not been tested and reported on in published work.

While cognitive readiness equips hoist operators to handle unpredictable hoist operations, the high-risk nature of these operations can impose significant psychological and emotional demands on the operator. Consequently, hoist operators must possess a high degree of psychological resilience and emotional regulation – another set of skills that is also absent from the core CRM framework. Recognising this gap, a specific psychological resilience factor was identified as an addition to the core non-technical skills, based on interviews with Italian HEMS crew, pilots, and hoist operators (Ceschi et al., 2019). The resulting NOTECHS+ self-report questionnaire may be useful for hoist trainee self-evaluation. Ceschi et al. (2019) added the categories of resilience and emotion regulation to the original NOTECHS conceptualisation and then tested and refined the new self-report survey using a sample of 211 participants recruited from 11 Italian HEMS bases. In this conceptualisation, resilience was defined as the capacity to safely adapt to changes and recover from stressful events, and emotion regulation as the automatic or intentional strategies used to manage emotions during stressful events. With both systems recognising skills beyond the core CRM skills, NOTECHS+ and the HeliNOTS systems are complementary; the HeliNOTS systems are directly observable behavioural markers (Hamlet et al., 2023), whereas the NOTECHS+ measures the trainee's self-evaluation of core non-technical, emotion regulation, and resilience skills (Ceschi et al., 2019). Both assessments could therefore be

incorporated into hoist operator training to measure trainee and instructor evaluation of hoist-relevant non-technical skills as training progresses.

Is VR simulation effective to train non-technical skills for hoisting?

Virtual reality simulation is a well-established tool for training non-technical skills, particularly in high-risk domains such as firefighting. Evidence from systematic reviews and meta-analyses indicates that VR is effective in inducing physiological stress and supporting experiential learning, often outperforming traditional training in knowledge acquisition and retention across a range of industries and contexts that include emergency response and firefighting, the military, aviation, and mining (Scorgie et al., 2024; Wheeler et al., 2021). Immersive VR scenarios incorporating unpredictable stressors have been shown to enhance decision-making, situational awareness, and communication under pressure during firefighter commander training (Wheeler et al., 2021). For example, Cohen-Hatton and Honey (2015) extended the standard operational training of firefighter commanders by adding three rapid mental checks (decision controls), which were then explicitly communicated to the broader team, prior to plan execution. The need for a focussed addition to standard training was identified following camera-based observation of live fire responses, and subsequent debriefing using the recorded visuals, in which the authors observed that commanders frequently jumped from information gathering to execution without formulating and evaluating a plan (Cohen-Hatton & Honey, 2015). Across their three studies, in which unexpected events that required a response were injected into the test scenarios, observed commander situational awareness was assessed on three levels: (1) perception of elements in the current situation; (2) comprehension of the current situation; (3) projection of future status (see Model of Situation Awareness e.g., Endsley (2015)). Cohen-Hatton & Honey (2015) demonstrated that by incorporating a pause and verbalising decision controls (Why am I doing this? What do I expect to happen? Are the benefits worth the risks?) into VR-based training improved subsequent plan formulation and anticipatory situational awareness measured through real-time audio/video recording of the live burn operational test scenarios, when compared to commanders who had received standard training only. This study highlights an important benefit of VR-based training – the ability for training scenarios to be manipulated through the safe incorporation of unexpected but realistic events to allow trainees to practice complex decision-making in risky scenarios. In similar manner, VR-based training can also improve emotional resilience and reduce negative affect, demonstrated through

stress induction training in military contexts (Pallavicini et al., 2016). Through repeated exposure to stressors, the stress response is reduced allowing trainees to better manage uncertainty and maintain high levels of performance under high levels of stress. In HEMS contexts, simulation-based training enables exposure to complex, costly, unsafe scenarios that are hard or impossible to replicate in live training without incurring risk to the crew and the aircraft, such as flying in low visibility or under challenging weather conditions (King et al., 2022; Pietsch et al., 2020; Sun et al., 2020). Exposing HEMS crew and hoist operators to these scenarios through VR-based training enables the trainees to learn how to effectively manage these events in a safe training environment, providing prior exposure rather than experiencing these situations for the first time in live missions.

Simulation-based training improves HEMS non-technical skills, including decision-making and communication, ideally these skills should be assessed using validated behavioural markers such as NOTECHS (O'Connor et al., 2008; Pietsch et al., 2016). Additionally, repeated VR training cycles have been initially linked to enhanced team effectiveness (Sun et al., 2020), although the proposed framework is yet to be empirically validated beyond the authors' initial evaluation. Another important advantage of simulation is its ability to increase training volume at low cost and risk. For example, indoor simulated environments allow significantly more hoist cycles per session than live training at substantially lower cost. These simulated environments have been successfully used to train HEMS crew for night operations, which is not trained for using live helicopters due to the level of risk (Pietsch et al., 2020). The increased practice frequency enabled by simulated training allows teams to refine communication patterns, develop cognitive strategies, and manage errors in a safe environment. Overall, the evidence supports high-fidelity, simulation-based training as an essential and efficient approach for developing the technical and non-technical skills required for safe and effective hoisting operations in complex, high-stakes environments. However, few studies aim to measure improvements in hoisting-related non-technical skills using simulated or VR-based training, particularly cognitive readiness and emotion regulation, and simulation fidelity is mostly untested in hoisting contexts.

Can VR Training Simulation enhance decision-making?

Recent research from the firefighting context demonstrates that expectations formed during simulated training can override explicit operational instructions, which may have consequences for hoist training. For instance, Tai et al. (2026), in semi-immersive VR SAR training, found that participants trained in high-hazard

environments committed more false-positive errors when later tested in a low-hazard environment. Notably, even when participants were explicitly instructed that the hazard prevalence had decreased, their self-generated expectations from prior training consistently drove their behaviour. The reliance on prior experience persisted under high-stress conditions (Tai et al., 2026). During live helicopter hoist operations, operators are constantly required to monitor for environmental threats, such as cable snag points, obstacles, or other environmental changes. Thus, there is a risk that an operator's decision-making and hazard perception may be biased by expectations formed during simulator scenarios, rather than adapting to the reality of the live mission. Because there is currently no empirical evidence in the HEMS or hoisting literature evaluating how these expectations might affect an operator's cognitive flexibility, it is essential that future VR training designs deliberately and randomly vary scenario parameters, for example, the accuracy of pre-training mission briefings, to prevent the development of expectation biases.

A concern raised when relying on VR-simulation for high-stakes emergency training is the presumption that the absence of real-world consequences might inadvertently lead to overconfidence and promote riskier decision-making than would occur in live operations (Arias et al., 2021). This assumption stems from the fundamental nature of simulators, which intentionally remove physical danger. Researchers caution that without the threat of actual harm or cost trainees might make decisions that significantly differ from how they would behave in a real-world scenario (Arias et al., 2021). For example, Russell et al. (2024) found that participants using an immersive VR headset during a residential fire scenario were significantly more likely to take the riskier action of investigating the fire compared to those using a non-immersive desktop VR. However, as the authors highlight, without comparative baseline data from real-world fire incidents, it is difficult to determine whether these actions are indicative of increased risk-taking because there are no real-world consequences, or whether they in fact are a realistic representation of human behaviour due to the increased level of immersion (Russell et al., 2024). Contrasting evidence identified in this scoping review finds that when simulation achieves a high level of affective and psychological fidelity, it induces stress and psychological caution that prompt trainees to behave much as they would in reality (Harris et al., 2020). Corroborating this finding, Wheeler et al. (2021) reported that decision-making in firefighting VR scenarios was very similar to decision-making in live incidents and live burns in real-world operations.

We speculate that fundamental differences in the populations tested in the two sets of evidence contributes to the contrasting findings. When exposed to high-fidelity VR

simulations of novel experiences, the general public (Arias et al., 2021; Russell et al., 2024) may respond very differently to professionals (Harris et al., 2020; Wheeler et al., 2021) who approach these simulations as occupational training tools. The decision-making and stress responses of professionals when being trained using high-fidelity VR simulations may more accurately mirror their real-world operational behaviours, whereas the general public may respond with sensation-seeking or riskier behaviours (Russell et al., 2024), potentially interpreting the situation as a game and not recognising that they should respond appropriately to the risks or danger presented. Future training designs must not simply presume that VR induces risky behaviour; instead, training programs should maximise psychological and affective fidelity to induce realistic, cautionary behaviour. Future research should address these literature gaps by: (1) empirically testing population differences in VR risk-taking, conducting cohort comparisons evaluating active professional operators' decision-making alongside naïve control groups, and (2) contrasting subjective trainee decisions made during a VR simulation against objective data such as visual scans of the environment, physiological stress measures, proximity to environmental hazards, conning and velocity adjustments. Additionally, and ideally, decision-making data from live operations should be incorporated into these evaluations, potentially through re-creating past missions in the simulator, to verify whether real-life decision-making mirrors live operational risk management. Together, these approaches would help to ensure that operators develop a grounded, realistic appreciation of hoisting operational risks.

How could we standardise and optimise VR hoist operator training?

If we consider that virtual reality training can be beneficial for hoist operators, the next question is how the VR training should be optimised. To optimise training design for hoist operators, training pathways must balance progressive difficulty with cognitive load management to effectively train both technical and non-technical skills. While not standardised across all jurisdictions, European industry guidelines emphasise that comprehensive hoist training must systematically combine technical task rehearsal with explicit non-technical skill development (Ceschi et al., 2019; EASA, 2024; Hamlet et al., 2020).

In general, initial training pathways should be carefully considered, as exposing novices immediately to complex scenarios can hinder learning. Scaffolded learning approaches should be utilised, which involves pitching each training mission just above the trainee's current level and providing supports that make it achievable

(Vygotsky, 1978). These supports are later withdrawn as competence develops. For example, Valente et al. (2025) demonstrated that using a sandtable-based trial-and-error approach significantly reduced trainees' cognitive load, resulting in higher test scores and faster completion times compared to traditional, high-demand training formats. Indoor HEMS simulation pathways apply this principle by scaffolding trainees' learning through hierarchically mapped, progressive scenarios, which move from daytime procedural baselines over safe terrain to complex nighttime winching in extreme weather and multi-system emergencies (Pietsch et al., 2020; Sun et al., 2020). By coupling each progressive scenario phase with an instructor-led debrief, trainees can consolidate procedural schemas before facing higher environmental demands (Pietsch et al., 2020).

Consistent with cognitive load theory's emphasis on managing intrinsic and extraneous load to protect capacity for schema construction, reducing difficulty through scaffolded learning is most beneficial for tasks with highly interacting elements (Pyke et al., 2025; Sweller, 1994), such as hoist operations where physical cable management, communicative, and monitoring demands must be processed together. In contrast, an alternate approach to training design is based on the desirable difficulty principle, in which training that is slower, more effortful, and produces more errors (impairing performance) has been shown to enhance long-term retention or transfer (Bjork & Bjork, 2011). This increased difficulty may be more beneficial for simpler tasks with fewer interacting elements, such as procedural tasks (Pyke et al., 2025). Pyke et al.'s (2025) integrative account therefore implies that training pathways should incorporate both theoretical approaches, depending on the complexity and interactive nature of the training mission, the experience of the hoist operator, and the planned learning outcomes. The need for scaffolding is most pronounced for novice hoist operators, who can build a robust foundation through lower cognitive load tasks before being progressively exposed to increasingly complex events, with scaffolding systematically faded as training progresses. Experienced hoist operators, however, are expected to have developed the schemas needed to manage the highly interacting elements of hoisting operations, so their training should increase cognitive load to match their higher capacity.

Standardised training pathways and tailoring to individual differences could be combined if standardisation is applied at the level of the framework rather than the individual training pathway. One approach would be to define a small number of standardised trainee categories, for example based on hoisting experience (e.g., novice, intermediate, and advanced hoist operators), each with standardised guidelines specifying scenario progression, the supports provided and how these are

faded out, and the competency criteria for progression. Competency could be measured, for example, using the behavioural marker systems described above (see Ceschi et al., 2019; Flin, 2010; Hamlet et al., 2023). The fading out of supports is consistent with evidence that instructional supports that benefit novices can become redundant, or even detrimental, for more experienced learners (e.g., Rey & Buchwald, 2011). Adaptation within each category, such as adjusting difficulty in response to workload indicators, could follow standardised decision rules, so that the process of tailoring is consistent across trainees and instructors even where the content delivered is not. Currently, hoisting experience is the only characteristic with preliminary support (e.g., King et al., 2022). While the literature highlights what types of content should be included, there is no evidence on how to optimally schedule training delivery for hoist operators. The review did not identify any evidence evaluating broader design factors, such as whether spaced versus massed learning schedules yield better real-world transfer for hoisting operations.

To operationalise scaffolding within MR hoisting training, instructional design can employ structured cognitive pauses and decision controls to manage cognitive processing during complex scenarios. Empirical evidence demonstrates that segmenting immersive VR lessons and embedding structured cognitive pauses significantly improves learning retention and transfer by preventing working memory overload (Parong & Mayer, 2018). When paired with goal-oriented decision prompts, such as the decision controls described above, requiring trainees to explicitly verbalise their rationale significantly improves plan formulation, preventing unsafe, reflexive decision-making (Cohen-Hatton & Honey, 2015) or channelled attention. In the context of helicopter hoist operations, cognitive pauses and decision controls translate directly, because live hoisting demands simultaneous physical cable management, aircraft conning, and obstacle monitoring under time pressure. Trainee hoist operators are vulnerable to channelled attention during unexpected events, as demonstrated in the 2011 crash example above (Plamp, 2011), where a focus on troubleshooting inside the cabin obscured aircraft drift and led to a cable snag. By embedding structured cognitive pauses during VR simulation at critical mission points, such as initial hover positioning, technical malfunctions, or escalating weather hazards, instructors can halt continuous motion to include explicit decision controls.

Efficient visual search or field scanning is another skill that can be scaffolded and improved using gaze training within VR simulations (Harris et al., 2021; Vine et al., 2014). For example, in the medical domain, an expert's visual gaze is overlaid into the VR visual field, and trainees are instructed to align their eyes and attention with these expert visual anchors, to avoid scanning inefficiently or becoming distracted by

background elements (Vine et al., 2014). Similar paradigms have been used in VR-based SAR hazard tasks to ensure trainees identify threats, or as an evaluative measure of core task focus (Tai et al., 2026; Walls et al., 2024). As hoisting operations require the hoist operator to monitor the environment for hazards, such as cable snag points, and may require a visual search to identify the location of the person to be rescued, scaffolded visual search training could be incorporated into training pathways for novice hoist operators. Scaffolding field scanning in MR simulations can help to reduce extraneous cognitive load as the ability to ignore less relevant information is improved, and to improve focus and attention on the core task.

Scaffolding need not be fixed in advance. In their recent systematic review of technical and experiential VR features, Tusher et al. (2024) argue that VR-based simulation training scenarios should be calibrated to the trainee's needs. One way this can be achieved is by adjusting training missions based on real-time feedback on the trainee's cognitive workload (e.g., Besson et al., 2013). To date, however, most VR-training evaluations rely on subjective questionnaires for stress and workload, with few evaluating objective measures such as physiological data, and they tend to focus on skills or knowledge development (Tusher et al., 2024). Despite the general lack of empirical focus on non-technical skills, studies have demonstrated that immersive VR successfully engages users at a physiological level, with objective metrics like heart rate and eye-tracking revealing significantly higher engagement and time-on-task compared to desktop simulations (Walls et al., 2024). From an individual calibration perspective, Besson et al. (2013) used heart rate variability (HRV) and skin conductance to build real-time cognitive workload estimation models for helicopter pilots in flight simulators – these types of models could be investigated for use during MR training sessions. Recreating physiological arousal is, in principle, the most accurate way to ensure a simulation achieves true physiologic fidelity (Evans et al., 2023), and by actively monitoring metrics like HRV to identify exactly when a hoist operator's cognitive capacity is challenged, real-time workload models could be used to dynamically adjust the scenario's variability and difficulty to match the trainee's stress tolerance. However, there can be measurement issues with physiological data collection in immersive simulation studies (e.g., Besson et al., 2013; Clifford et al., 2019). These are likely to be particularly relevant to hoist operator training, in which physical exertion might be expected when the operator is required to exert physical control over the cable during, for example, higher wind conditions. Physical exertion makes changes in heart rate associated with cognitive workload difficult to disentangle from changes caused by physical workload (Mattingsdal et al., 2025), and other confounds,

including postural changes, respiratory patterns and the clothing worn in training, such as flight suits, may make many biomarkers impractical or ineffective for assessing stress during immersive training scenarios (Clifford et al., 2019; Tai et al., 2026). Identifying objective markers of cognitive workload and stress in immersive simulation and real-world scenarios is therefore a pressing issue.

Given the practical challenges and signal noise associated with capturing physiological responses during active, physical simulations, future training designs could investigate less invasive, objective measures of cognitive overload and stress, such as visual fixations and acoustic voice analysis (Besson et al., 2013; Clifford et al., 2019; Mattingsdal et al., 2025; Tai et al., 2026). The reviewed literature establishes that eye-tracking metrics provide insights into a user's attention and cognitive load, with longer visual fixations implying deep cognitive processing, and may indicate (sometimes inappropriately) channelled attention (Harris et al., 2020; Walls et al., 2024). Promisingly, the current Air Rescue training system has already begun to capture where a trainee is looking and to display this gaze data on the instructor's impact screen. In the hoisting context, instructors could therefore use this gaze tracking to immediately identify if a trainee is experiencing cognitive overload, evidenced by fixating too long on a single distractor or failing to maintain a broad visual scan of the environment (Harris et al., 2020; Walls et al., 2024). As another alternative or complement, analysing recorded verbal communications for vocal stress markers can objectively track trainee psychological arousal. Stress can change a speaker's voice, as stress stiffens the laryngeal muscles and vocal cords (Veiga et al., 2025). Both physical exertion and stress would be expected to make changes to the acoustic profile of a trainee's voice; however, these have not been directly compared in the literature to date. Features such as loudness, which may increase with physical effort and decrease as a result of stress, speech rate and breathing-related pauses – plus information from the simulator indicating when the trainee is exerting physical effort over the cable, for example – could be investigated as ways to distinguish between then (Anikin, 2023; Paulmann et al., 2016). Voice analysis could therefore be used to detect stress-induced voice changes during communication, which could be used to measure and track, for example, improved resilience over time.

By integrating these types of physiological measures, instructors gain access to real-time, objective insights into a hoist operator's psychological state without the physical encumbrance and measurement errors associated with wearable physiological sensors (Besson et al., 2013; Tai et al., 2026). Incorporating these alternative metrics into the VR-based hoisting training may therefore allow instructors

to objectively monitor a trainee's situational awareness and stress levels, and to adjust scenarios so that they are calibrated to the trainee's cognitive capacity and not too demanding for effective learning to take place (Tusher et al., 2024; Valente et al., 2025). These data could also be fed back to trainees. In their review of methods for teaching non-technical skills in emergency medicine training, Brevik et al. (2023) advocate for using physiological feedback as an active teaching strategy, such as pairing video review of gaze data with post-training mission debriefs (extending the instructor-led debriefs described above), or even incorporating, for example, heart rate data into the VR display, so that trainees can recognise cognitive overload and freezing behaviours, and practise self-regulation techniques. Capturing these measures consistently across trainees and instructors could also provide the objective progression criteria needed for a standardised pathway, although this remains to be tested in the hoisting context.

Can hoist training be optimised to account for trainees' characteristics?

While absent in the hoisting and HEMS simulation literature, evidence from other domains demonstrates that individual characteristics predict VR learning outcomes. For example, prior knowledge, motivation, and spatial abilities influence VR-based skills learning (Tusher et al., 2024), and selective attention and executive control influenced VR-based science learning, but working memory capacity did not (Lawson & Mayer, 2024). In the medical domain, participants with a sociable, stable emotional profiles exhibited decreased motivation and lowest acceptance of VR, whereas less sociable participants exhibited the opposite (Mühling et al., 2025); the authors suggested that the (often) solitary nature of VR training appeals to some personalities more than others. However, participants who reported a mix of emotions towards VR, but higher motivation and generally positive VR acceptance, complicated this simple sociability contrast. Additionally, the participants who responded most positively towards VR had higher technical affinity and prior VR experience. The influences of personality and prior technology exposure and how these factors support or oppose an individuals' VR acceptance are therefore difficult to disentangle (Mühling et al., 2025). Given these complex interactions and the current absence of empirical research examining learner characteristics within the hoisting and HEMS literature, future research should profile the individual cognitive, psychological, and experience-based characteristics of hoist trainees to better understand how to optimise VR hoisting training.

Conclusion

This scoping review evaluated the current evidence regarding VR simulation for helicopter hoist operator training, focusing on its operational value, underlying psychological mechanisms, optimal design, and individual predictors of learning. Hoist operator training is not standardised across jurisdictions, and there are no industry-wide guidelines on the skills or proficiency levels required. The broader literature indicates that VR and MR simulations offer a safe, cost-effective alternative or complement to live training, enabling repeated exposure to high-risk or rare scenarios without endangering the crew or the aircraft. MR systems incorporating haptic feedback make it feasible to measure hoist-specific technical skills, and emerging evidence suggests that experienced and novice operators can be distinguished by how they complete training missions. Evidence on non-technical skills training, however, comes largely from other high-risk domains, and there is no consensus on which non-technical skills hoist operators require beyond core CRM skills.

Optimising MR-based training design requires deliberate scaffolding of training pathways to prevent cognitive overload, particularly early in training. Cognitive pauses in which trainees verbalise their plans may build habits that support decision-making as missions become more complex and emergency scenarios are introduced, while scaffolded gaze training may support efficient environment scanning and visual search. For experienced hoist operators, skills currency and emergency-scenario training should be tailored to their higher capacity. Instructors could calibrate difficulty for each trainee using voice acoustics and gaze direction, which may be more practical than physiological measures. Psychological and affective fidelity appear more consequential than physical realism; maximising them may encourage the realistic, cautious decision-making expected in live operations. Scenario parameters should also be varied to prevent trainees forming expectations that do not hold in live missions.

Overall, the literature on VR and MR training for hoisting is in its infancy. The foundational value of simulation for skills acquisition and stress induction is well supported, but there is no evidence yet of transfer to live missions, no trainee evaluations of simulator fidelity, no evidence on optimal delivery schedules (for example, MR only versus combined live and MR training, or spaced versus massed practice), and little evidence on which individual characteristics predict hoisting ability or learning trajectories. Addressing these gaps forms the basis of the Phase 2 priorities set out in Chapter 3.

Chapter 2. VR Training Evaluation – Testing the Benefits and Limitations for Hoist Operations

Air Rescue produced and implemented, with design support from the project team, a survey to measure trainee perceptions of the VR simulations. The survey was designed so that it could measure trainees' experience across their training pathways, after key training missions and at the end of a training session to measure perception of the last training mission as well as their perception of VR training overall. Developed while the scoping review was in progress, the survey primarily as a record of the training mission characteristics and serves as a subjective self-report tool for

- cognitive load, perceived training utility and perceived learning of each training mission
- system usability (ergonomic, physical fidelity) at the end of the training session.

Survey Development

Measures

Perceived Learning. Trainees were asked to evaluate the benefits of their most recent training mission on a Likert scale from 1 – No not at all to 4 - Yes, very much. Five questions asked if the trainee had learned something new, found the training mission to be useful, whether the training mission helped them to refine their hoisting ability, and, separately, their knowledge, and whether the training mission had stretched their skills.

Subjective Workload. The survey adapted the NASA Task Load Index (NASA-TLX: Hart & Staveland, 1988) to measure trainees' subject workload of each training mission. Trainees rated their experience of their most recent training mission across

six dimensions on a slider scale from 0 to 100. The six dimensions measure: mental demand - the amount of mental and perceptual activity required; physical demand - the amount of physical activity required; temporal demand - the level of time pressure felt; effort - how hard the trainee had to work to achieve their performance; performance - how successful the trainee felt they were in performing the task; and frustration - how irritated, stressed, or annoyed versus content or relaxed the trainee felt. Two descriptors were provided above each slider scale. For mental, temporal, physical demand and effort these were very low and very high, displayed above each end of the scale; for performance the descriptors were failure and perfect, and content/relaxed and stressed/irritated for frustration. The dimension scores were averaged to create a raw global workload score, reversing the performance scale as this scale is reversed in the original conceptualisation (Hart & Staveland, 1988). The NASA-TLX, as originally specified, also surveys each trainee's evaluation the importance or contribution of each dimension, in addition to dimension workload scores. These additional data are using to create a set of which are then applied to the individual dimension scores to produce a weighted global workload score (Hart & Staveland, 1988). The survey did not collect these weightings data; this could be considered in future evaluations of the hoisting training.

Training Mission Evaluation. Trainees were asked three questions to evaluate their perceptions of the training mission they had just experienced. Trainees evaluated the difference between the training mission they had just experienced and their prior experiences in either a simulator or in real-life operations. Responses were recorded on a Likert scale from 1 – No, not at all to 3 – Yes, completely. Trainees also rated their perception of the training mission duration and difficulty on slider scales from 0 to 100, with three descriptors displayed above the scale (duration: too short, just right, and too long; difficulty: too easy, just right, too difficult).

Training mission Description. Trainees reported up to 7 training mission descriptions (multiple answers were accepted) that characterised their most recent training mission. These descriptions were: emergencies, vessel hoisting, wet hoisting, land open, land cliff/mountain, stretchers, and hi line/tag line.

If trainees indicated that they were completing the survey following their final training mission of the day, they then completed an additional set of questions that asked them to evaluate their overall experiences of the VR training mission simulations.

VR Quality. Trainees rated the quality of their training mission experience and the scene details on a 1 – 4 Likert scale (1 – Poor/bad, 2 – Fair, 3 – Good, 4 – Excellent).

VR Experiences. Trainees rated possible negative experiences of the VR simulation on a 1 – 5 Likert scale (1 – Not at all to 5 - A lot). In this block, trainees were asked whether they experienced any discomfort, fatigue, dizziness, or motion sickness. Trainees also rated the comfort of the headset on a 1 – 5 Likert scale (1 – Not at all comfortable to 5 - Very comfortable).

VR Evaluation. Trainees responded to five questions in which they reported their acceptance of the VR training missions as part of the hoisting training curriculum. Trainees reported the extent to which they agreed or disagreed that the VR training was valuable and, separately, acceptable for use in the standard hoisting curriculum, could reduce the amount of live training required, could replace some of the classroom-based hoisting training, and was a useful complement to live hoisting training. Responses were recorded on a 1- 5 Likert scale (1 – Strongly disagree to 5 – Strongly agree). The trainee's evaluation responses were supplemented with two open-ended questions that asked what the most valuable aspect of the VR hoisting training was, and what should be improved before using the training more widely.

Demographics

Hoisting Experience. Trainees recorded their level and type(s) of hoisting operational experience. Hoisting experience level was categorised as: 0-1 Years, 1-3 Year, 3–7 Years, 7-12 Years, 12-15 Years, and 15+ Years. Trainees could record up to 5 types of experience: Little to no experience, HEMS HHO, Wind farm/pilot transfer, O&G SAR/Lim SAR, All Weather SAR.

Participants

Air Rescue UK collected feedback from 13 trainees who had completed between 1 to 7 training missions in one training day – a total of 48 post-training mission evaluations. Of these, 3 trainees had provided feedback following 6 or 7 training missions in total, 3 trainees provided feedback after 4 or 5 training missions, 4 trainees had completed 3 training missions, and the remainder had recorded feedback after completing 1 or 2 training missions in the simulator. Of those trainees who recorded details of their hoisting experience, 4 trainees had between 0 and 1 years of experience, 2 trainees had 3 – 7 years hoisting experience, and 3 had 15 or more years hoisting experience. Examining the type of experience the hoist operators reported, 4 trainees recorded that they had limited or no hoisting experience, 3 trainees had experience in SAR only, and 2 trainees had experience bridging HEMS, wind farms or pilot transfers, and SAR operations. Experience data was not collected for the remaining 4 trainees.

Table 1. Hoist trainees who took part in the training evaluation and their characteristics

Experience Level	Number of individuals	End of training mission form	End of training form
Novice (0 -1 years)	4	3	6
Intermediate (3 - 7 years)	2	2	3
Advanced (15+ Years)	3	7	6
Mixed (Unknown)	5	18	3

Preliminary Results

The following summarises the data collected between May 26th and July 2nd, 2026. The survey was actively developed during the data collection period; therefore, not all items were responded to by all trainees. Where there are missing data for data items, this is highlighted.

Perceived Learning

As shown in Figure 1, most trainees were positive in their evaluation of learning after each VR hoisting training mission. Of the 25 responses. 84 – 92% of post-training mission evaluations indicated the trainee had learned at least a little during their VR training mission, and in 44-48% of these evaluations the trainees felt “yes, very much” that the training mission was useful, exposed them to something new, helped them refine their knowledge and ability, and stretched their skills.

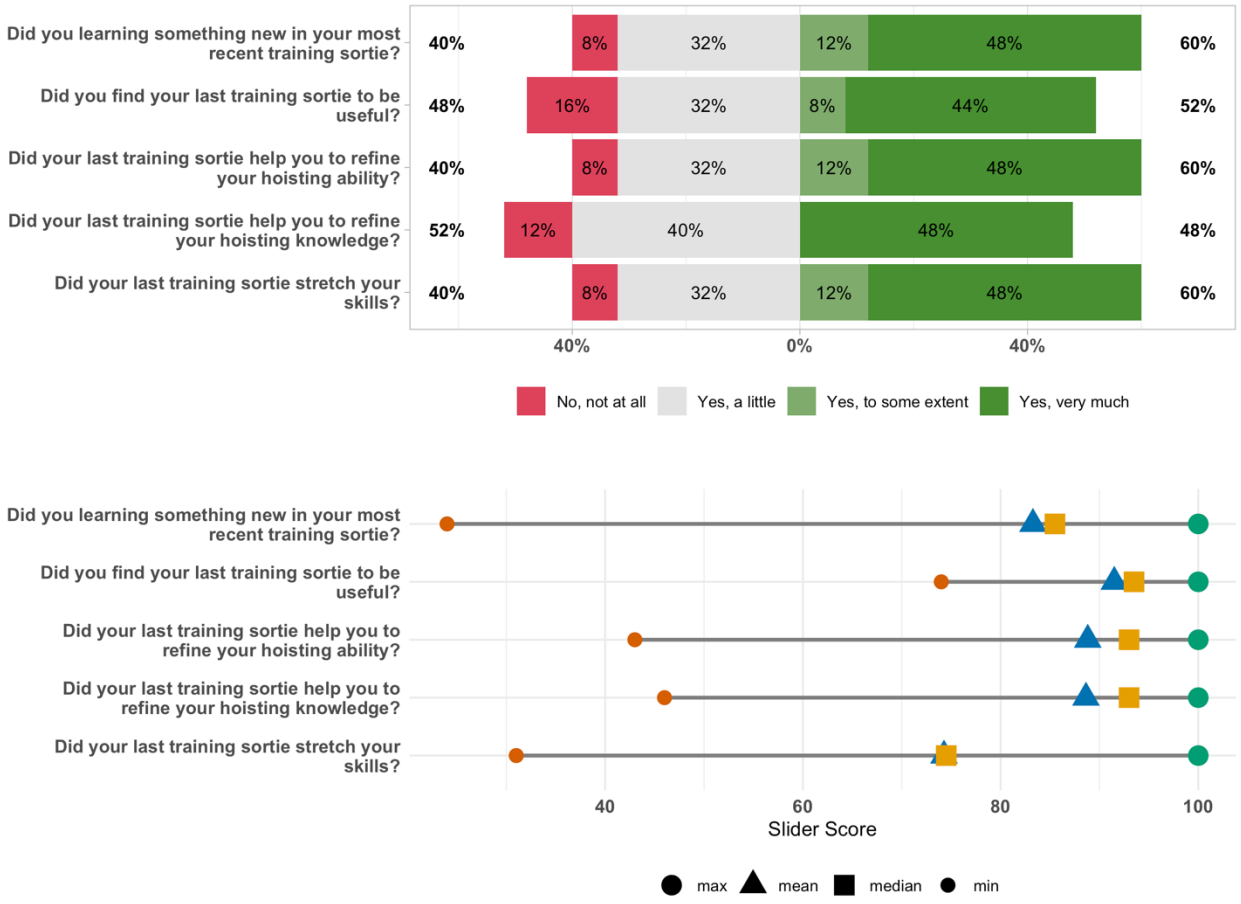


Figure 1. Evaluation of learning by trainees after training mission completed. Panel A – evaluated on a Likert scale (*n* = 25 responses). Panel B: evaluated on a 0 – 100 slider scale (*n* = 20 responses).

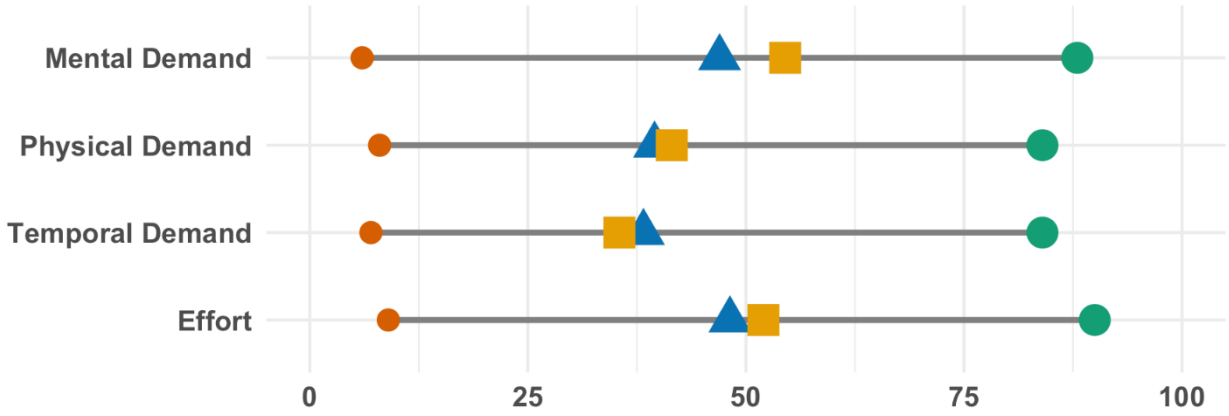
The data collection for learning effects was changed part-way through the data collection period. A further 20 post-training mission responses were completed using a 0 -100 slider scale for each question. Average scores for each question ranged from 67% (stretching skills) to 92% (usefulness). In general, the pattern of responses suggested that most of the responses were very positive for learning something new and finding hoisting ability and knowledge.

Minimum scores suggested that there may be some responders who felt that their training mission experience was less novel. Together, these data suggest that most trainees report at least some learning in each of the training missions, but there may be some individuals who value the training missions less.

Subjective Workload

All trainees had completed an assessment of the subjective workload after selected training missions throughout their training day. There was wide variation in global workload evaluations reported by the hoisting trainees (*M* = 38.1, *SD* = 17.0, *Median* = 42.2, range: 6.8% to 71.5%), with 4 trainees reporting training mission global workload scores below 26% and 1 trainee reported a training mission score above 68%. As shown in Figure 2, average scores in NASA-TLX dimensions, indicated a high-level of perceived performance and low levels of frustration across training mission evaluations (performance: *M* = 73.2, *SD* = 20.3; frustration: *M* = 26.0, *SD* = 21.9). Perceived mental workload effort were close to 50% across training missions (mental: *M* = 47.0, *SD* = 24.5; effort: *M* = 48.2, *SD* = 25.1), perceived physical and temporal workload were similar to each other but lower (physical: *M* = 39.5, *SD* = 21.9; temporal: *M* = 38.2, *SD* = 23.2).

Compared to a meta-analysis of NASA-TLX scores by task types, global workload median and average scores across hoisting training missions were below the midpoint (50% of the scores / percentile) reported for command-and-control type tasks, and similar to midpoint for cognitive tasks, and aircraft pilot tasks (Grier, 2015). The range observed from these hoisting trainee reports are wider than (outside) the middle 80% of scores observed across all operational domains included in the meta-analysis (Grier, 2015).



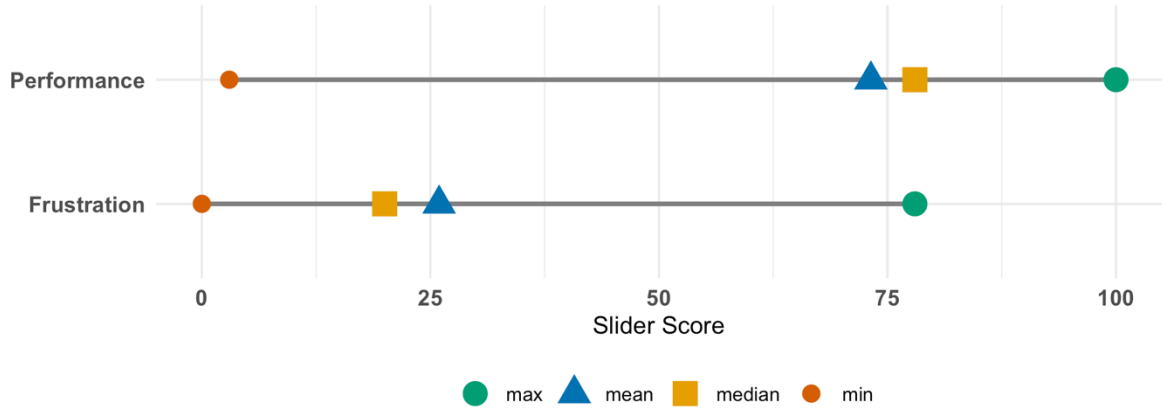


Figure 2. NASA-TLX subjective workload dimension scores reported by trainees across all post-training mission evaluations (*n* = 48).

Subjective workload may require some optimisation for certain types of trainees. For example, some trainees reported training mission performance scores of 100%, or close to 100%, (*n* = 4 trainees of unknown experience across 4 training missions), and very low frustration scores of 0%, or close to zero (*n* = 3 trainees of unknown experience). Minimum demand and effort scores are also low in some evaluations. The global minimum is lower than the minimum reported for most task types within the meta-analysis. We highlight, however, that NASA-TLX scores should be considered a guide only – indeed, the meta-analysis reported a general lack of interpretability of the NASA-TLX scores, and there are no guidelines on how to interpret optimum workload using this measure.

Training Mission Evaluation (novelty, duration, difficulty)

In addition to evaluating their learning from each training mission, the trainees reported their perception of the novelty, duration, and difficulty of each simulated training mission. Most trainees perceived that the training mission they had just completed had at least some novelty in comparison to their previous experienced, although a minority did not perceive any novelty in the training mission.

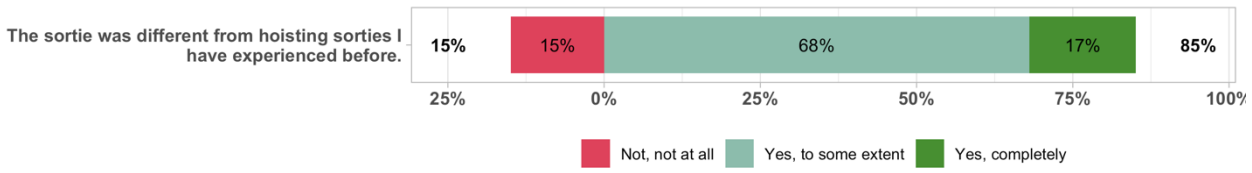


Figure 3. Trainees' perception of the novelty of the training mission they had just completed.

Across the training mission evaluations, trainees scored their training missions as "just right", however the range of responses suggest that there were some training missions that may have been perceived as tending towards being too long, and others as tending towards being too easy and too short.

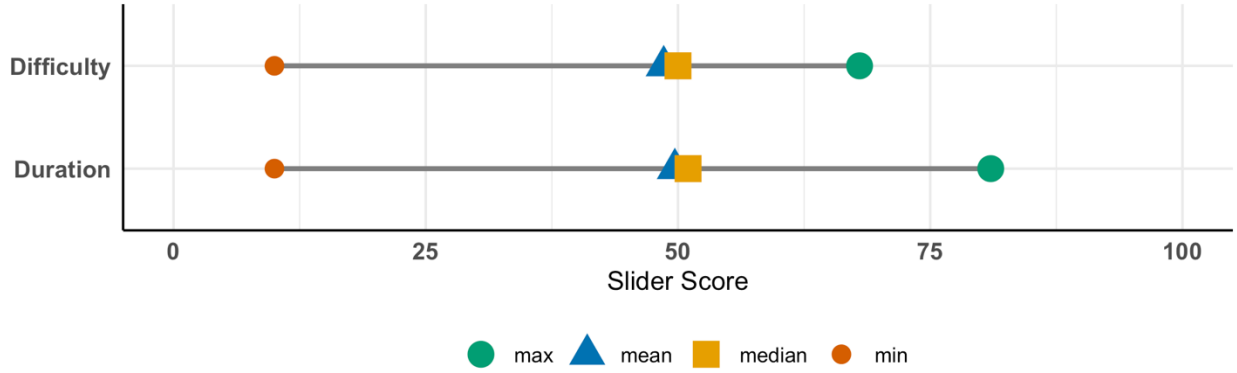


Figure 4. Trainee's perception of the difficulty and duration of the training mission they had just completed.

What factors predict learning?

Correlational analyses were used to examine whether NASA-TLX subjective workload, and perceived training mission characteristics (novelty, duration, and difficulty) were associated with learning outcomes. Across the 43 training mission evaluations, in which trainees had a response to each question on the post-training mission survey, the NASA-TLX global score was strongly correlated with perceived usefulness of the training mission, and the trainees' perceptions that they could refine their knowledge and ability and stretch their skills (see Figure 5). Associations between novelty and duration with learning were less clear.

Subgroup analyses of the two largest groupings by experience level – highly experienced hoist operators with 15+ years' experience compared with a mixed group of hoist operators for whom hoisting experience had not been recorded – revealed that associations differed depending on experience levels (see Figures 5 and 6). For example, while the correlation between difficulty and learning outcomes were very small and not significant across all training missions, the correlation coefficients were moderate to large, but in opposing directions for the mixed and experienced groups. For experienced hoist operators, duration was strongly and negatively associated with refining knowledge, and there was no correlation between duration and NASA global score. These data suggest that experienced hoist operators did not perceive the longer sorties to be more challenging, and that there

was little benefit to longer training missions for these individuals. However, for the mixed group the association was positive and significant for perceived usefulness, refining ability and knowledge and stretching skills; the correlation between duration and NASA global scores was not significant. The pattern for perceived difficulty opposed that for duration, with the mixed group perceiving less benefit from difficulty, whereas the experienced group strongly associated difficulty with novelty, refining their ability and knowledge, and stretching their skills. Despite this preference, perceived difficulty and global workload scores were not correlated for experienced hoist operators but were for the mixed group. Examining the underlying NASA-TLX dimensions, focussing on

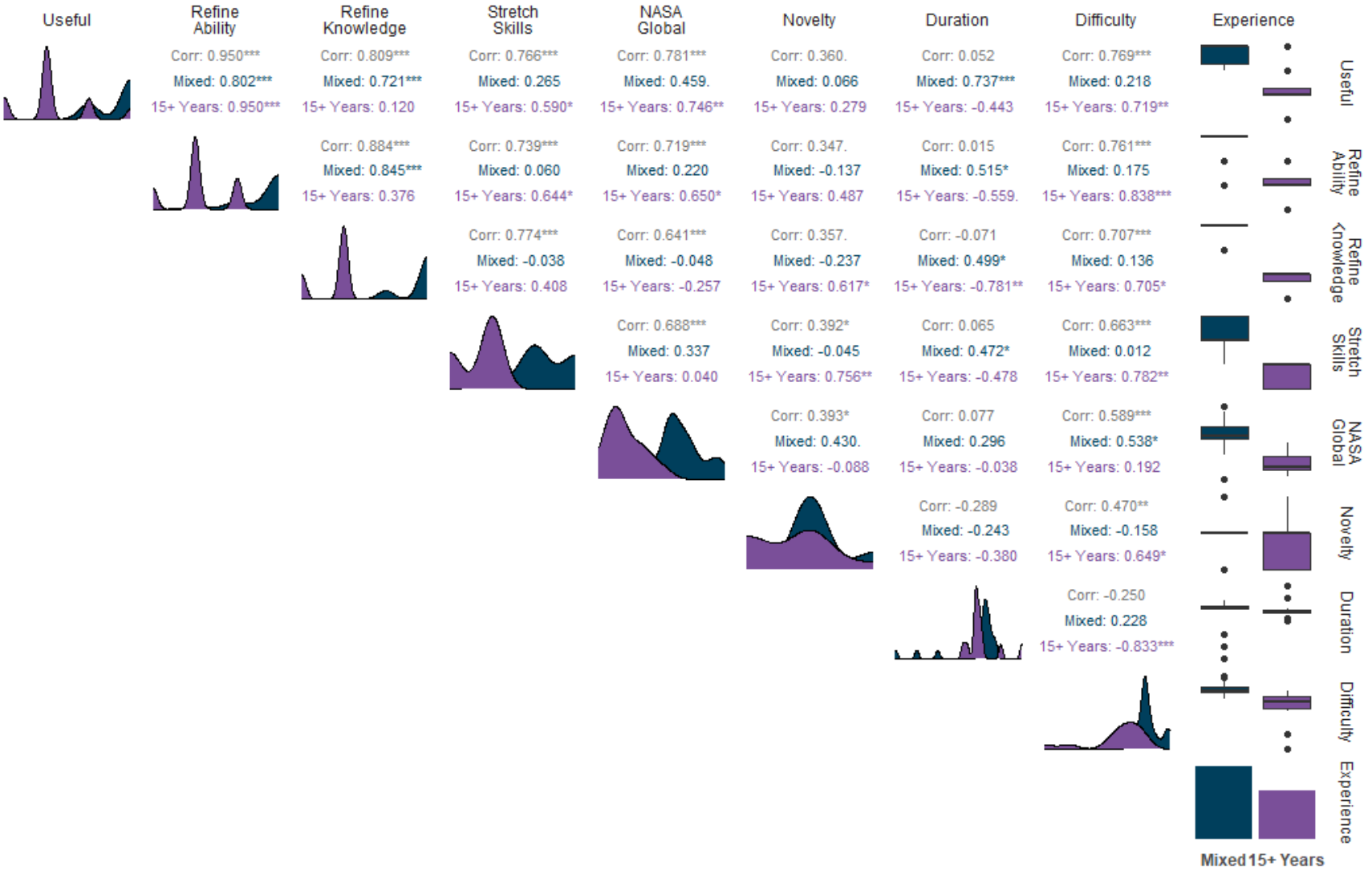


Figure 5. Correlations between the raw NASA-TLX global score and perceived novelty, duration, and difficultly with learning effects. Something new learning effect removed (perfect correlation).

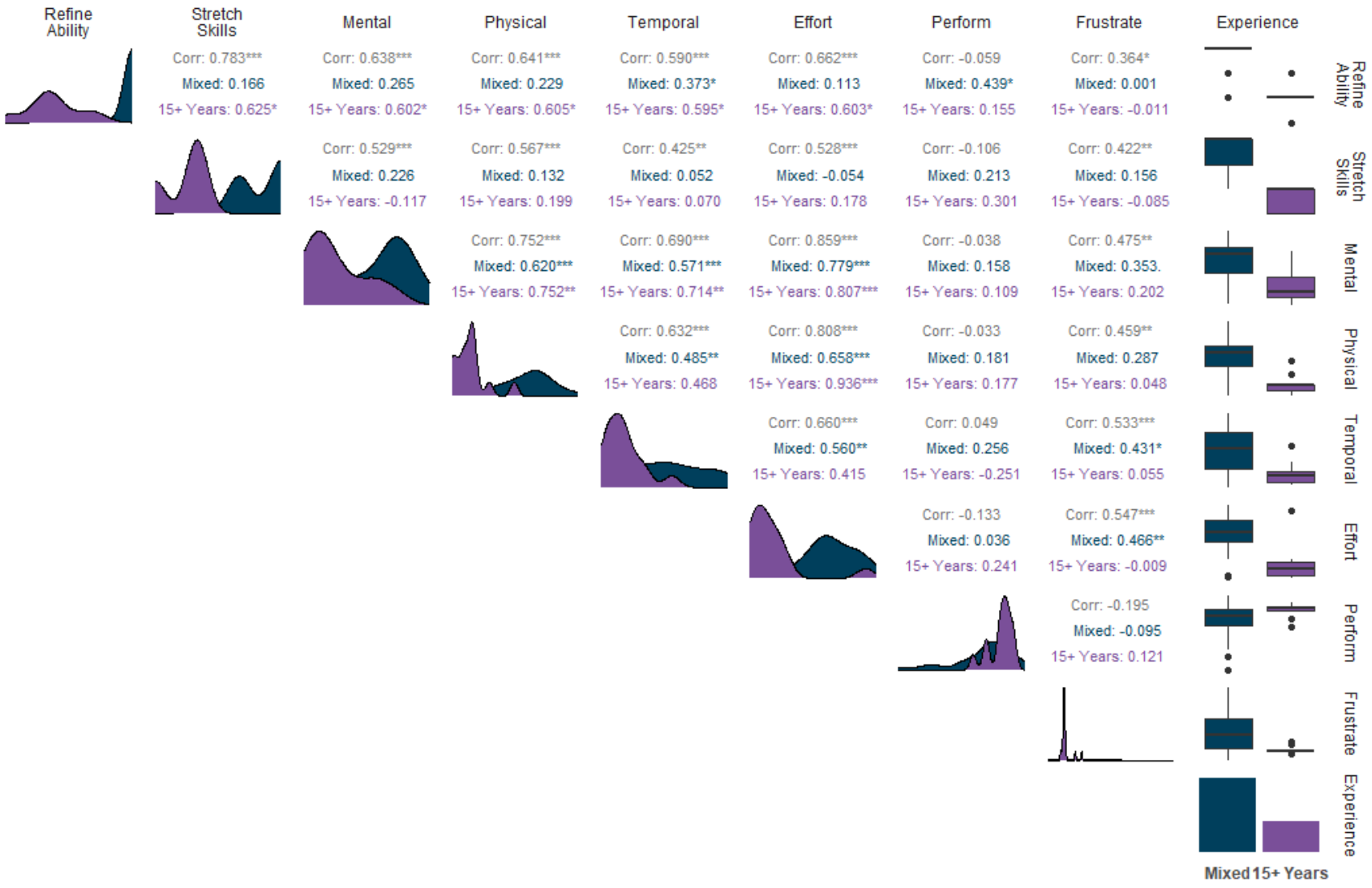


Figure 6. Correlations between the raw NASA-TLX dimension scores with refine ability and stretch skills learning effects.

trainees' perceptions of the training mission refining their ability or stretching their skills, suggested that the highly experienced hoist operators associated mentally challenging, higher effort and more pressured hoisting training missions with refining their ability and stretching their skills. The mixed group perceived that a higher level of performance and time pressure resulted in a significant, moderate associations with refining ability. There were no significant correlations between NASA dimensions and the perception that the training mission stretched their skills.

Overall, we conclude that the mixed set of results suggest there are group differences in the perceived NASA subjective workload dimensions, and that experience level changes perceptions of what is important for learning in the VR training missions. However, we also highlight that for the purposes of this preliminary data analysis, the between-groups discussions here are largely speculative at this stage. A larger dataset with more observations for each of the experience levels is required to disentangle exactly how individual NASA subjective workload dimensions and duration, novelty, and difficulty are associated with learning outcomes based on hoisting experience.

VR quality, comfort, and acceptability for hoist training

Following their last training mission of the training session, trainees evaluated their overall experiences of the VR hoisting simulations.

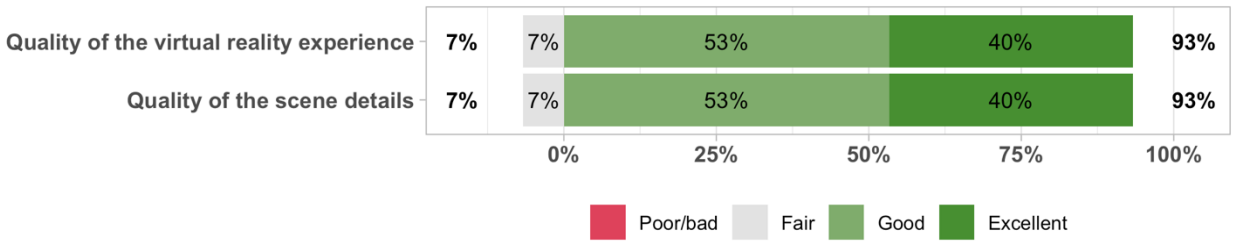
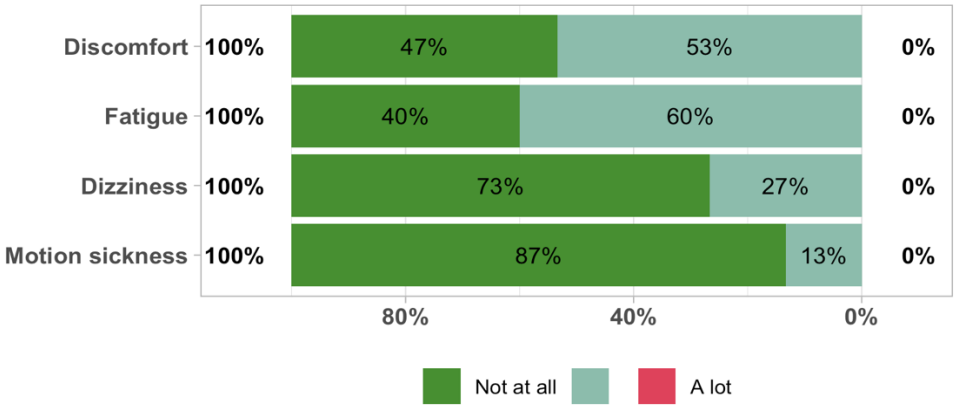


Figure 7. Trainee perceptions of the VR experience and scene quality.

VR Quality & Experiences. Trainees were very positive about the quality of the VR simulations, with almost all trainees rating the quality as good or excellent (see Figure 7). Most trainees did not experience any dizziness or motion sickness during their training missions, with a minority reporting a little of these (see Figure 8). A little fatigue was more likely to be reported – this may be linked to duration and/or difficulty which was higher for some participants.



Most trainees reported that they found the VR headsets to be quite or very comfortable.

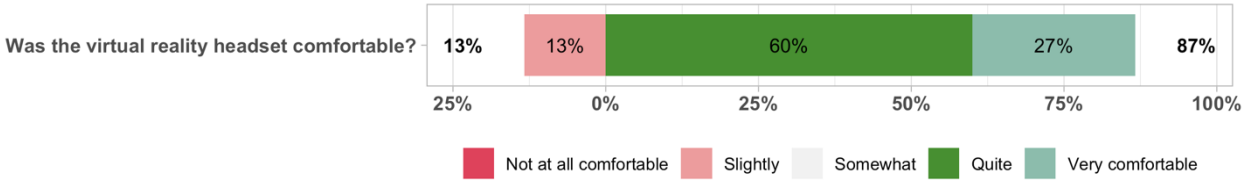


Figure 8. Headset comfort reported by the trainees.

VR value and acceptability for hoist training. Most trainees agreed or strongly agreed that the VR hoisting simulations were valuable and acceptable for use in the standard hoisting curriculums, and that this type of training was a useful complement to and could reduce the amount of, live hoisting training required (see Figure 9). Most trainees also agreed or strongly agreed that the VR hoisting training could replace some classroom-based training, although for this question, there were a small minority of participants who disagreed.

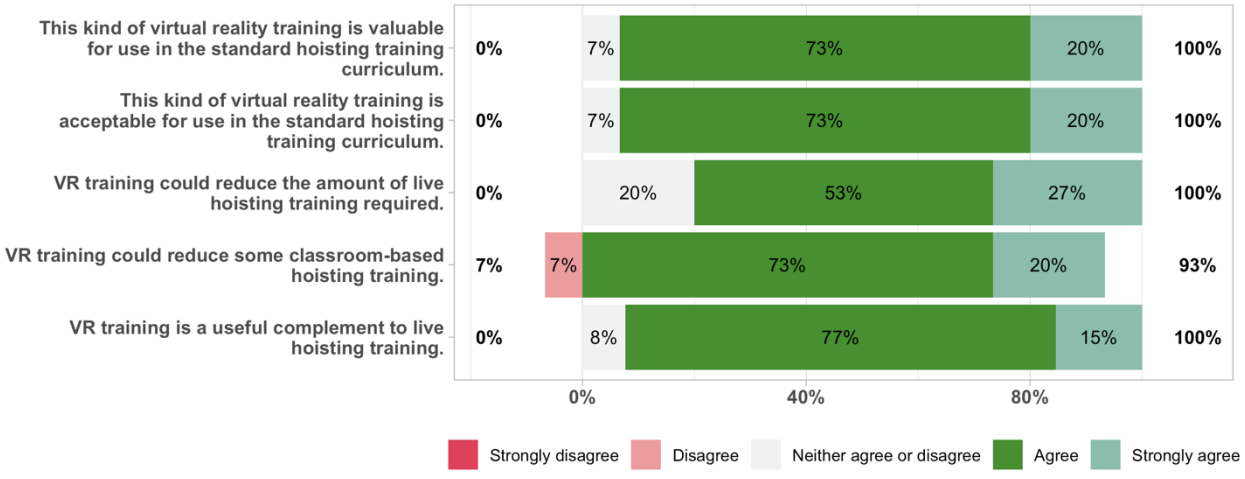


Figure 9. Trainee's overall evaluations of the VR hoisting simulations at the end of their training day.

Open-Response Questions: Perceived Value and Future Improvement.

To complement the quantitative survey data, trainees were able to provide further feedback through two open-response questions. When asked about the most valuable aspect of the VR training, responses overwhelmingly focused on the ability to safely execute high-risk, real-time emergency procedures that are either impossible or entirely too dangerous to replicate in actual flight. Trainees specifically praised the opportunity to practice critical reactions to events such as winch-runaways, cable-fouls, engine failures whilst winching, and simulated hoist cuts. Additionally, trainees highly valued the logistical efficiency of the virtual environment, noting the distinct advantage of running back-to-back training missions in difficult conditions without the operational constraints of refuelling or consuming live aircraft flight hours.

When asked what should be improved before the training is implemented more widely, the feedback almost exclusively centred on the need for greater physical and functional fidelity tailored to specific operational environments. Some trainees emphasised that the simulator's hardware and scenarios must perfectly mirror the distinct aircraft types they operate on. They called for exact replications of aircraft-specific equipment, controls, and fits, such as integrating a physical "real hoist cut button" into the operator's setup. Furthermore, trainees noted that if they are to use off-roster time for this training, the VR curriculum must seamlessly complement the specific procedures employed during their mandated live currency training, particularly for those operating on dual-hoist aircraft. Finally, while the system was generally well-tolerated, minor improvements to headset comfort were also suggested.

Conclusion

The intra-day survey conducted by Air Rescue UK captures trainees' perceptions of virtual reality simulations, demonstrated the training's value and suggested areas for design improvement. Overall, the preliminary data strongly support the value of MR hoisting simulations as a viable training tool. Trainees overwhelmingly agreed that the MR hoisting simulations are a valuable and acceptable addition to the hoisting curriculum, serving as a useful complement that could reduce the amount of high-risk live training required and even replace some classroom-based instruction. Furthermore, trainees reported substantial learning effects. Between 84% and 92% of post-training mission evaluations indicated that trainees learned at least a little, with almost half strongly indicating that the simulations were highly useful, exposed them to new scenarios, and effectively stretched their skills.

To evaluate the optimal design and psychological effects of the training, the survey measured subjective cognitive workload. . The highest workload scores were not overly high (Grier, 2015), however several more experienced hoist operators reported global workload scores that may be lower than optimal, close to minimum subjective load evaluations reported for air traffic control tasks, and below those for command-and-control and cognitive tasks. As more trainee evaluations are collected, future work could examine more closely which training missions these lower evaluations relate to and consider whether this is only related to hoist operator experience, specific to some, or results from a combination of these two factors.

While most trainees perceived the overall novelty, duration, and difficulty of the training missions as optimal, responses varied significantly. Some trainees reported very low subjective workload, effort, frustration and novelty, suggesting certain training missions might not be taxing enough, while others found the scenarios to be too long or too difficult. The correlational analysis shows a significant relationship between the global subjective workload scores a trainee experiences during a training mission and the educational benefit they perceive. Novelty scores significantly correlated with the trainee's perception that the training mission had stretched their skills and helped refine their knowledge for more experienced hoist operators. While not significant, the associations were negative for the mixed experience group. The link between novelty and training outcomes may also differ between experience-levels and be more important for experienced hoist operators. Novelty may be a motivational and enjoyment factor for less experienced trainees (Huang et al., 2024), but may relate to the nature of the training mission itself for experienced hoist operators. It is possible that the general scenario itself does not need novel e.g., a hoisting training mission over landscape new to the trainee, but that a combined set of factors (e.g., weather conditions, physical access restrictions, visual constraints) may be needed to make a training mission more novel and stretch the skills of more experienced hoist operators. Raising novelty through a combination of factors may also lead to higher subjective workload scores for experienced hoist operators. Because learning and ability refinement are closely tied to the level of subjective workload, the data highlight that a tailored approach to training may be optimal.

To maximise training efficacy, instructors should consider dynamically calibrating the difficulty and novelty of the VR scenarios to the individual user, particularly based on their prior hoisting experience. For instance, highly experienced operators reported the greatest learning benefits from shorter, highly mentally challenging, and time-pressured emergency scenarios. Finally, the survey confirmed that the MR system is highly rated by trainees. Trainees highly rated the visual quality of the MR experiences and found the headsets to be quite or very comfortable. The immersive simulations were well tolerated; most trainees experienced no dizziness or motion sickness, though a minority reported mild fatigue, which may be associated with differing perceived task durations and difficulty. Qualitative feedback highlighted that while operators valued the ability to safely practice high-risk emergencies like cable-fouls and winch-runaways, future iterations of the simulator could consider the benefits of exact physical and functional fidelity, including aircraft-specific controls and tactile features like a real hoist cut button. The evidence from the scoping review suggests that improving physical realism and face validity is less important for training outcomes, but may be important for trainee acceptance (Harris et al., 2020). It should be considered that these types of suggestions for improvement may be closely related to lower subjective workload evaluations – increasing subjective workload evaluations may be a more effective response to this feedback than increasing further simulator physical and face fidelity. Future work should also consider whether language barriers may influence how the trainees respond to the survey questions.

Chapter 3. Recommendations for Phase 2+ Research

Based on the literature gaps identified in the scoping review and the qualitative and quantitative findings from the preliminary survey, the following non-exhaustive recommendations are made for future research. These recommendations build directly on the literature gaps and survey findings and are intended to advance the overall project goals of demonstrating training value, identifying psychological mechanisms, optimising training design, and predicting learning outcomes. Feasibility, prioritisation, and phasing of these recommendations have not been established at this stage. In addition to these recommendations, investigating optimal training delivery scheduling would also be highly beneficial, but may be best addressed in a future phase given it may need a partner organisation to compare live-only vs combined VR-based and live training, and how best to schedule initial training pathways for hoist operators.

1. Validate Fidelity, Construct Validity, and Real-World Skill Transfer

The scoping review identified a lack of empirical evidence confirming that technical and non-technical skills acquired in mixed-reality hoisting simulations, and from non-VR simulated scenarios, successfully transfer to live-flight missions. Qualitative feedback from the Air Rescue UK survey emphasised the need for improved physical and functional fidelity, with trainees requesting additional tactile features – although at this stage it is unclear if these changes would result in improved learning outcomes, or improved user acceptance, or perhaps both (Harris et al., 2020). It is important, therefore, to systematically assess the simulator's fidelity using established frameworks, such as the one proposed by Harris et al. (2020), as this is not currently established in the existing literature. Existing construct validity work in this domain has compared experienced military hoist operators against general-population volunteers rather than trainee hoist operators; future validation studies should use comparison groups drawn from trainee populations to ensure findings generalise to the population Air Rescue UK is training. Beyond this, comparative studies tracking whether trainees trained in the VR simulator demonstrate measurable performance improvements in real-world, live-flight hoisting operations compared to traditional training cohorts would be beneficial. While addressing this gap will further support the business case for replacing some live-flight training with mixed-reality hoisting training, it should be noted that the wider literature suggests that there is a strong business case through cost avoidance (e.g., Pietsch et al., 2020), and particularly, the avoidance of risk for emergency-type scenarios (e.g., Scorgie et al., 2024). Demonstrating real-world transfer of skills would provide further weight to argument.

2. Implement Reliable, Objective Measures of Cognitive Load and Stress

While measuring physiological arousal is, in principle, the most accurate way to assess affective fidelity, the review noted that wearable sensors (such as heart rate monitors) suffer from signal noise due to the physical exertion – which is required during active hoisting – and can be further impacted by wearing flight suits, commonly worn during professional training sessions. Rather than replacing physiological measurement outright, Phase 2 could prioritise validating less invasive, objective metrics for stress and cognitive load against physiological measures in a subsample where physical exertion confounds can be modelled and controlled, to establish which combination of measures provides the most reliable signal in the active hoisting context. Specifically, future research should investigate using the gaze-tracking data already captured on the instructor's impact screen to evaluate situational awareness and cognitive processing (e.g., analysing the length and frequency of visual fixations). Additionally, acoustic voice analysis could be employed to detect stress-induced pitch elevations during crew communication.

3. Develop Adaptive, Experience-Tailored Training Pathways

The preliminary survey results indicated that the training mission evaluations differed by trainees' hoisting experience levels, with learning outcomes and perceived workload varying accordingly. Phase 2 could expand data collection beyond the current 13 trainees to profile the specific training mission factors (duration, novelty, difficulty, type of hoist operation) that are driving higher cognitive loads for less experienced trainees, and which factors can be added to existing scenarios to stretch the skills of more experienced hoist operators. Dynamic training models could then be tested, where the instructor uses real-time metrics to actively calibrate the difficulty and novelty of the training missions to optimally match the specific stress tolerance and experience level of the individual learner.

4. Profile Individual Psychological Predictors of VR Learning

Beyond hoisting experience, evidence from other domains indicates that individual characteristics such as prior VR technology exposure, spatial ability, personality, and motivation predict VR learning outcomes (Lawson & Mayer, 2024; Mühling et al., 2025; Tusher et al., 2024). This evidence base is absent from the HEMS and hoisting literature. Phase 2 could incorporate validated cognitive and psychological measures (e.g., spatial ability, attention, technical affinity) alongside the existing training mission-evaluation survey, to establish which individual differences – beyond experience level – predict learning trajectories and VR acceptance for hoist trainees.

5. Validate Specific Non-Technical Skills Assessment Frameworks

The review highlighted the need for standardised assessment of hoisting non-technical skills (such as cognitive readiness, adaptability, and resilience) but noted that current behavioural marker systems are largely untested in MR hoisting environments. Phase 2 could systematically test and integrate HEMS-specific frameworks into the MR hoisting training. This includes testing the HeliNOTS (SAR and O) systems for instructor-based observation of behavioural markers, and the NOTECHS+ scale for trainee self-evaluation of emotion regulation and resilience. Building on the broader emergency literature, this phase could explicitly incorporate the active teaching of managing psychological arousal and adapting to changing situations.

References

Abich, J., Parker, J., Murphy, J. S., & Eudy, M. (2021). A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality*, *25*(4), 919–933. https://doi.org/10.1007/s10055-020-00498-8

Abrahamsen, H. B., Sollid, S. J., Öhlund, L. S., Røislien, J., & Bondevik, G. T. (2015). Simulation-based training and assessment of non-technical skills in the Norwegian Helicopter Emergency Medical Services: a cross-sectional survey. *Emergency Medical Journal*, *32*(8), 647–653. https://doi.org/10.1136/emermed-2014-203962

Anikin, A. (2023). The honest sound of physical effort. *PeerJ*, *11*, e14944.

Arias, S., Wahlqvist, J., Nilsson, D., Ronchi, E., & Frantzich, H. (2021). Pursuing behavioral realism in Virtual Reality for fire evacuation research. *Fire and materials*, *45*(4), 462–472.

Ash, J. S., & Smallman, C. (2008). Rescue Missions and Risk Management: Highly Reliable or Over Committed? *Journal of Contingencies and Crisis Management*, *16*(1), 37–52. https://doi.org/https://doi.org/10.1111/j.1468-5973.2008.00528.x

Besson, P., Bourdin, C., Bringoux, L., Dousset, E., Maïano, C., Marqueste, T., Mestre, D. R., Gaetan, S., Baudry, J. P., & Vercher, J. L. (2013). Effectiveness of Physiological and Psychological Features to Estimate Helicopter Pilots' Workload: A Bayesian Network Approach. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, *14*(4), 1872–1881. https://doi.org/10.1109/TITS.2013.2269679

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, *2*(59-68), 56–64.

Braddee, R. W., Washenitz, F. C., II, Mezzanotte, T. P., Romano, N. T., Pettit, T. A., Merinar, T. R., & McDowell, T. W. (2000). Six Career Fire Fighters Killed in Cold-Storage and Warehouse Building Fire – Massachusetts. https://doi.org/https://doi.org/10.26616/NIOSHFFFACE99F47 (Fire Fighter Fatality Investigation and Prevention Program (FFFIPP))

Brevik, C., Miller, D., Kendall, J., & Michael, S. (2023). Nontechnically speaking: A review of tools and methods in the teaching and assessment of nontechnical skills in emergency medicine training. *AEM Education and Training*, *7*(6), e10911. https://doi.org/10.1002/aet2.10911

Ceschi, A., Costantini, A., Zagarese, V., Avi, E., & Sartori, R. (2019). The NOTECHS+: A Short Scale Designed for Assessing the Non-technical Skills (and more) in the Aviation and the Emergency Personnel. *Front Psychol*, *10*, 902. https://doi.org/10.3389/fpsyg.2019.00902

Clifford, R. M. S., Jung, S., Hoermann, S., Billinghurst, M., Lindeman, R. W., & Ieee. (2019). *Creating a Stressful Decision Making Environment for Aerial Firefighter Training in Virtual Reality* 2019 26TH IEEE CONFERENCE ON VIRTUAL REALITY AND 3D USER INTERFACES (VR),

Cohen-Hatton, S. R., & Honey, R. C. (2015). Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. *Journal of Experimental Psychology: Applied*, *21*(4), 395–406. https://doi.org/10.1037/xap0000061

Endsley, M. R. (2015). Situation Awareness Misconceptions and Misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4–32. https://doi.org/10.1177/1555343415572631

European Union Aviation Safety Agency. (2024). *ESPN-R Helicopter Hoist Pilot Training Guide*. https://www.easa.europa.eu/en/downloads/140785/en

Evans, A. J., Russo, C. M., Tovar, M. A., Liu, A., & Conley, S. P. (2023). Physiologic Fidelity as a Domain in Assessing Mixed Reality Trauma Simulation. *MILITARY MEDICINE*, *188*(11-12), 3322–3329. https://doi.org/10.1093/milmed/usac244

Flin, R. (2010). Chapter 6 - CRM (Non-Technical) Skills — Applications for and Beyond the Flight Deck. In B. G. Kanki, R. L. Helmreich, & J. Anca (Eds.), *Crew Resource Management (Second Edition)* (pp. 181–202). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-374946-8.10006-8

Gordon, M., Darbyshire, D., & Baker, P. (2012). Non-technical skills training to enhance patient safety: a systematic review. *Med Educ*, *46*(11), 1042–1054. https://doi.org/10.1111/j.1365-2923.2012.04343.x

Grier, R. A. (2015, 2015). How high is high? A meta-analysis of NASA-TLX global workload scores. Human Factors and Ergonomics Society Annual Meeting, Los Angeles, CA.

Hamlet, O. E. D., Irwin, A., Flin, R., & Sedlar, N. (2023). The development of the helicopter non-technical skills (HeliNOTS) behavioural marker systems. *Ergonomics*, *66*(12), 2232–2241. https://doi.org/10.1080/00140139.2023.2194591

Hamlet, O. E. D., Irwin, A., & McGregor, M. (2020). Is It All about the Mission? Comparing Non-technical Skills across Offshore Transport and Search and Rescue Helicopter Pilots. *The International Journal of Aerospace Psychology*, *30*(3-4), 215–235. https://doi.org/10.1080/24721840.2020.1803746

Harris, D. J., Bird, J. M., Smart, P. A., Wilson, M. R., & Vine, S. J. (2020). A Framework for the Testing and Validation of Simulated Environments in Experimentation and Training [Review]. *Frontiers in Psychology*, *Volume 11 - 2020*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.00605

Harris, D. J., Hardcastle, K. J., Wilson, M. R., & Vine, S. J. (2021). Assessing the learning and transfer of gaze behaviours in immersive virtual reality. *Virtual Reality*, *25*(4), 961–973. https://doi.org/10.1007/s10055-021-00501-w

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight Simulator Training Effectiveness: A Meta-Analysis. *Military Psychology*, *4*(2), 63–74. https://doi.org/10.1207/s15327876mp0402_1

Huang, X., Zhao, Q., Liu, Y., Harris, D., & Shawler, M. (2024). Learning in an Immersive VR Environment: Role of Learner Characteristics and Relations Between Learning and Psychological Outcomes. *Journal of Educational Technology Systems*, *53*(1), 3–29. https://doi.org/10.1177/00472395231216943

Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2021). The Effects of Virtual Reality, Augmented Reality, and Mixed Reality as Training Enhancement Methods: A Meta-Analysis. *Hum Factors*, *63*(4), 706–726. https://doi.org/10.1177/0018720820904229

King, M., Lenser, S., Rogers, D., & Carnahan, H. (2022). Novice and experienced hoist operators in a helicopter hoist virtual reality simulator. *International Journal of Training Research*, *20*(1), 1–13. https://doi.org/10.1080/14480220.2021.1933569

King, M. T., Lenser, S., Rogers, D., & Carnahan, H. (2020). Helicopter hoist performance in novice and experienced hoist operators. *Aerospace Medicine and Human Performance*, *91*(6), 496–500.

Lawson, A. P., & Mayer, R. E. (2024). Individual differences in executive function affect learning with immersive virtual reality. *Journal of Computer Assisted Learning*, *40*(3), 1068–1082. https://doi.org/10.1111/jcal.12925

Mattingsdal, H., Dahle, T. N., Abrahamsen, E. B., Søvik, S., & Ottestad, W. (2025). Workload in helicopter rescue operations – A comparison of two different rescue methods in a randomized cross-over design. *SAFETY SCIENCE*, *192*, 106996. https://doi.org/10.1016/j.ssci.2025.106996

Mühling, T., Backhaus, J., Demmler, L., & König, S. (2025). How Personality and Affective Responses Are Associated with Skepticism Towards Virtual Reality in Medical Training—A Pre-Post Intervention Study [Article]. *Cyberpsychology, Behavior, and Social Networking*, *28*(5), 335–341. https://doi.org/10.1089/cyber.2024.0567

O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E., & Wilson, K. A. (2008). Crew Resource Management Training Effectiveness: A Meta-Analysis and Some Critical Needs. *The International Journal of Aviation Psychology*, *18*(4), 353–368. https://doi.org/10.1080/10508410802347044

O'Neil, H. F., Lang, J., Perez, R. S., Escalante, D., & Fox, F. S. (2014). What Is Cognitive Readiness? In H. F. O'Neil, R. S. Perez, & E. L. Baker (Eds.), *Teaching and Measuring Cognitive Readiness* (pp. 3–23). Springer US. https://doi.org/10.1007/978-1-4614-7579-8_1

Pallavicini, F., Argenton, L., Toniazzi, N., Aceti, L., & Mantovani, F. (2016). Virtual Reality Applications for Stress Management Training in the Military. *Aerosp Med Hum Perform*, *87*(12), 1021–1030. https://doi.org/10.3357/amhp.4596.2016

Parong, J., & Mayer, R. E. (2018). Learning Science in Immersive Virtual Reality. *Journal of Educational Psychology*, *110*, 785–797.

Paulmann, S., Furnes, D., Bøkenes, A. M., & Cozzolino, P. J. (2016). How Psychological Stress Affects Emotional Prosody. *PLOS ONE*, *11*(11), e0165022. https://doi.org/10.1371/journal.pone.0165022

Pietsch, U., Knapp, J., Mann, M., Meuli, L., Lischke, V., Tissi, M., Sollid, S., Rauch, S., Wenzel, V., Becker, S., & Albrecht, R. (2021). Incidence and challenges of helicopter emergency medical service (HEMS) rescue missions with helicopter hoist operations: analysis of 11,228 daytime and nighttime missions in Switzerland. *Scandinavian journal of trauma, resuscitation and emergency medicine*, *29*(1), 92. https://doi.org/10.1186/s13049-021-00898-y

Pietsch, U., Knapp, J., Ney, L., Berner, A., & Lischke, V. (2016). Simulation-Based Training in Mountain Helicopter Emergency Medical Service: A Multidisciplinary Team Training Concept. *Air medical journal*, *35*(5), 301–304. https://doi.org/10.1016/j.amj.2016.05.006

Pietsch, U., Lischke, V., Sollid, S. J. M., Prückner, S., Theiler, L., Greif, R., & Albrecht, R. (2020). Efficacy and efficiency of indoor nighttime human external cargo mission simulation in a high-fidelity training Centre. *Scandinavian journal of trauma, resuscitation and emergency medicine*, *28*(1), 61. https://doi.org/10.1186/s13049-020-00755-4

Plamp, C. C. (2011). *USAF UH-1N SAR Helicopter Hoist Training Accident* (Air Accident Investigation Board Report, Issue. https://aerossurance.com/helicopters/usaf-hoist-training-accident/

Pyke, W., Lunau, J., & Javadi, A.-H. (2025). Does difficulty moderate learning? A comparative analysis of the desirable difficulties framework and cognitive load theory. *Quarterly Journal of Experimental Psychology*, *78*(10), 2181–2195.

Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: cognitive load and motivational explanations. *J Exp Psychol Appl*, *17*(1), 33–48. https://doi.org/10.1037/a0022243

Russell, M. D., Bonny, J. W., & Reed, R. (2024). Impact of Virtual Reality on Decision-Making and Risk Assessment During Simulated Residential Fire Scenarios. *Fire*, *7*(12), 427.

Scorgie, D., Feng, Z., Paes, D., Parisi, F., Yiu, T. W., & Lovreglio, R. (2024). Virtual reality for safety training: A systematic literature review and meta-analysis. *SAFETY SCIENCE*, *171*. https://doi.org/10.1016/j.ssci.2023.106372

Somerville, A., Joiner, K., Lynar, T., & Wild, G. (2025). Applications of extended reality in pilot flight simulator training: a systematic review with meta-analysis. *Visual Computing for Industry, Biomedicine, and Art*, *8*(1), 25. https://doi.org/10.1186/s42492-025-00206-w

Stefan, H., Mortimer, M., & Horan, B. (2023). Evaluating the effectiveness of virtual reality for safety-relevant training: a systematic review. *Virtual Reality*, *27*(4), 2839–2869. https://doi.org/10.1007/s10055-023-00843-7

Sun, X., Liu, H., Tian, Y. L., Wu, G. H., & Gao, Y. (2020). Team effectiveness evaluation and virtual reality scenario mapping model for helicopter emergency rescue. *CHINESE JOURNAL OF AERONAUTICS*, *33*(12), 3306–3317. https://doi.org/10.1016/j.cja.2020.06.003

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/https://doi.org/10.1016/0959-4752(94)90003-5

Tai, Y. S., Grange, J. A., & Honey, R. C. (2026). Self-Generated Expectations of Hazard Prevalence Affect Virtual Search and Rescue. *HUMAN FACTORS*, *68*(5), 580–598. https://doi.org/10.1177/00187208251410492

Tissington, P., & Flin, R. (2005). Assessing risk in dynamic situations: lessons from fire service operations. *Risk Management*, *7*(4), 43–51.

Tusher, H. M., Mallam, S., & Nazir, S. (2024). A Systematic Review of Virtual Reality Features for Skill Training. *Technology, Knowledge and Learning*, *29*(2), 843–878. https://doi.org/10.1007/s10758-023-09713-2

Valente, L., De Lorenzis, F., Calandra, D., & Lamberti, F. (2025). Enhancing Sand-Table-Based Incident Command Training With Extended Reality and Interactive Simulations: A Use Case in Forest Firefighting. *IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES*, *18*, 273–288. https://doi.org/10.1109/TLT.2025.3545436

Veiga, D. L., Almeida, T. M., Uchida, R. R., & Cordeiro, Q. (2025). The Fundamental Frequency of Voice as a Potential Stress Biomarker: A Systematic Review and Meta-Analysis. *Stress Health*, *41*(5), e70112. https://doi.org/10.1002/smi.70112

Vine, S. J., Moore, L. J., & Wilson, M. R. (2014). Quiet eye training: The acquisition, refinement and resilient performance of targeting skills. *European Journal of Sport Science*, *14*(S1), S235–S242. https://doi.org/https://doi.org/10.1080/17461391.2012.683815

Vygotsky, L. S. (1978). *Mind in Society*

Development of Higher Psychological Processes. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Walls, R., Nageswaran, P., Cowell, A., Sehgal, T., White, T., McVeigh, J., Staykov, S., Basett, P., Mitelpunkt, D., & Sam, A. H. (2024). Virtual reality as an engaging and enjoyable method for delivering emergency clinical simulation training: a prospective, interventional study of medical undergraduates [Article]. *BMC Medicine*, *22*(1), Article 222. https://doi.org/10.1186/s12916-024-03433-9

Wheeler, S. G., Engelbrecht, H., & Hoermann, S. (2021). Human Factors Research in Immersive Virtual Reality Firefighter Training: A Systematic Review [Review]. *Frontiers in Virtual Reality*, *2*, Article 671664. https://doi.org/10.3389/frvir.2021.671664