International Journal of Industrial Ergonomics 107 (2025) 103741

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

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

Beyond certification: Improving lifeguard drowning detection through validated tools and specialized training

Benjamin T. Sharpe[a,*], Marcus S. Smith[b,c], Steven C.R. Williams[c], Adam Hampshire[c,d], Maria Balaet[d], William Trender[d], Peter J. Hellyer[c,d], Jo Talbot[e], Jenny Smith[b]

[a] Institute of Psychology, Business, and Human Sciences, University of Chichester, Chichester, UK
[b] Institute of Applied Sciences, University of Chichester, Chichester, UK
[c] Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[d] Department of Medicine, Imperial College London, London, UK
[e] Royal Life Saving Society (RLSS), UK

ARTICLE INFO

Keywords:
Drowning detection
Intervention
Lifeguard
Memory
Vigilance

ABSTRACT

This study investigated two key aims: (1) the external validity of an animated performance assessment tool previously utilized in lifeguard training, with a focus on how lifeguard experience and task duration affect performance metrics, and (2) the impact of two distinct training protocols on lifeguard-specific drowning detection abilities. In the first experiment, experienced lifeguards demonstrated superior performance compared to inexperienced lifeguards in both 30-min tasks; however, both groups exhibited a decline in performance over time. The external validity of the animated tool was supported by its ability to produce performance outcomes aligned with real-world lifeguard tasks. The second experiment revealed that training specifically designed for lifeguard drowning detection significantly enhanced detection performance, while working memory training showed no measurable effect. These results highlight the necessity of incorporating realistic drowning detection challenges—such as varied bather numbers, drowning durations, and locations—into lifeguard certification programs, which currently do not emphasize these critical elements. The study also points to the significant proportion of lifeguards who missed drowning scenarios at baseline, underscoring the urgent need for improved training. Future research should explore the potential of animated tools in training and further investigate the cognitive mechanisms that underpin effective drowning detection.

1. Introduction

A lifeguard's capacity to remain attentive towards the unlikely event of a drowning scenario unfolding appears to be limited (Sharpe et al., 2023; Sharpe and Smith, 2024). Empirical evidence for such limitation is particularly notable in aquatic environments where the number of bathers fluctuate (Lanagan-Leitzel, 2021; Vansteenkiste et al., 2021) and task lengths appear to extend human capacity (Sharpe et al., 2024). Irrespective of such challenges, those with greater lifeguarding experience do outperform their lesser experienced counterparts (Lanagan-Leitzel and Moore, 2010; Laxton et al., 2021; Page et al., 2011). However, it is often noted that lifeguards who demonstrate this higher level of performance appear to hold an unexplored mechanism that facilitate sustained attention across extended periods (Laxton et al., 2022). This is more evident given that even the most experienced lifeguards do not independently detect all drowning scenarios across extended periods (Sharpe et al., 2023, 2024).

1.1. Sustained attention and working memory capacity

The ability to continuously maintain attention to a task, often referred to as sustained attention, is a core aspect of attention control. Sustained attention plays a crucial role across all occupational environments and is increasingly recognized as a key area of focus in human factors and ergonomics research (e.g., Bao et al., 2024; Chen et al., 2022; Foroughi et al., 2023; Greenlee et al., 2024; Pak et al., 2024). Attention control, commonly discussed in terms of Norman and Shallice's (1986) conception of the supervisory attentional system, allows an individual to direct attention towards stimuli to ensure the attainment of the current goal and minimize external distraction (Corbetta and Shulman, 2002;

* Corresponding author. Institute of Psychology, Business, and Human Sciences, University of Chichester, Chichester, PO19 6PE, UK.
E-mail address: B.Sharpe@chi.ac.uk (B.T. Sharpe).

https://doi.org/10.1016/j.ergon.2025.103741
Received 25 September 2024; Received in revised form 27 February 2025; Accepted 2 April 2025
Available online 9 April 2025
0169-8141/

Petersen and Posner, 2012). Sustained attention enables the maintenance and engagement on a task for extended periods (Robertson and Garavan, 2010), whilst a failure of such ability inevitably leads to a vigilance decrement (i.e., the decline in performance during extended monitoring tasks; Warm and Parasuraman, 1987). The vigilance decrement is thought to occur due to resource depletion over time, as proposed by attentional resources theory (Warm et al., 2008). According to this perspective, maintaining vigilance is cognitively demanding and depletes limited attentional resources, leading to decreased performance over time. An alternative explanation is the mindlessness theory, which suggests that the monotony of vigilance tasks leads to task disengagement (Robertson et al., 1997). Both accounts help explain why detection performance tends to deteriorate during prolonged surveillance tasks such as lifeguarding.

Working memory capacity (WMC) is considered a mechanism responsible for retaining a small amount of information in an active state for use in ongoing tasks (Baddeley, 2007; Miyake and Shah, 1999). Prior reports have demonstrated that WMC is significantly and positively correlated with multiple attention control measures (Unsworth and Robison, 2017). Reports have suggested that higher-WMC individuals are better able to sustain their attention than their lower cognitively advantaged counterparts (Buehner et al., 2006; Schweizer and Moosebrugger, 2004; Unsworth and Robison, 2017, 2020). Unsworth and Robison (2020) reported that those with high or low cognitive ability initially perform similarly in sustained attention tasks; however, as task duration increased, those with low WMC experienced a greater vigilance decrement than those with higher WMC. This relationship between WMC and sustained attention can be explained by the executive attention theory of working memory (Engle and Kane, 2004), which proposes that individuals with higher WMC can better maintain goal-relevant information in the face of distraction. In the context of lifeguarding, this suggests that experienced lifeguards may have developed knowledge structures and recognition strategies that enable more efficient allocation of attentional resources, reducing cognitive load and enabling them to sustain attention for longer periods.

Recent preliminary findings suggest cognitive abilities may systematically discriminate lifeguard drowning detection performance where lifeguards with greater WMC outperform lower WMC counterparts in an animated drowning detection task (Sharpe et al., 2024). Findings demonstrated that lifeguards with higher WMC were better able to maintain a greater level of performance over time, whilst those with lower WMC saw the greatest vigilance decrement. Results appear to support the contention that a higher degree of WMC may enable an individual to better maintain cognitive control and avoid periodic failures in goal maintenance (Engle and Kane, 2004). The detection of drowning events shares similarities with change detection paradigms studied in cognitive psychology. Change detection involves identifying modifications in visual scenes, which can be particularly challenging when changes occur gradually or during interruptions (Rensink, 2002). Similarly, drowning detection requires identifying critical changes in swimming behaviour amidst numerous distractors. Research has demonstrated that expertise influences change detection performance through enhanced selective attention and pattern recognition (Gorman et al., 2018), suggesting that experienced lifeguards may possess superior change detection abilities specific to drowning behaviours.

1.2. Cognitive training

A means to potentially enhance functions that relate to drowning detection performance may take two routes, including formal detection training (e.g., Laxton et al., 2023) and cognitive training (e.g., Dehnabaei et al., 2024). Both options appear to have conflicting findings within the literature; however, no such investigation has been demonstrated with respect to lifeguard drowning detection. Cognitive training has received substantial interest with mixed results (e.g., Shipstead et al., 2012). Although WM training has been demonstrated to improve additional cognitive mechanisms (e.g., fluid intelligence; Jaeggi et al., 2008, 2012), a proportion of research has failed to demonstrate far transfer effects beyond the trained function (Ball et al., 2002; Brehmer et al., 2012; Dahlin et al., 2009; Redick et al., 2020). However, improvements to daily functions have been reported previously in healthy participants (Ball et al., 2007; Basak et al., 2008; Edwards et al., 2009; Jaeggi et al., 2011; Karbach and Kray, 2009; Willis et al., 2006). Literature suggests healthy participants may more reliably show cognitive plasticity associated with training (Fernandez-Ballesteros et al., 2012).

Utilizing a computerized adaptive n-back task (Jaeggi et al., 2008), Owen et al. (2013) explored whether WM training could increase the neural and behavioural filtering of relevant and irrelevant information in individuals using a change detection task. In line with a lifeguard's role to inhibit task-irrelevant information (i.e., discern between relevant and irrelevant information while supervising bathers; Schwebel et al., 2011), this initial evidence demonstrated improvements to attentional control (i.e., inhibitory function) following WM training (Owens et al., 2013). Attentional control processes are proposed to modulate individual differences in WMC (Engle and Kane, 2004; Kane et al., 2007; Kane and Engle, 2002), and that WMC predicted performance in a lifeguard-specific detection task (Sharpe et al., 2024), it may be warranted to explore whether WM training would elicit far transfer effects to a real-world drowning detection task.

1.3. Lifeguard detection training

Prior literature has explored the use of detection training interventions through discrimination tasks across various domains, such as airport baggage surveillance and military surveillance (Schuster et al., 2013; Guznov et al., 2017). While the utility of training perceptual abilities has seen mixed results (Abernethy and Wood, 2001), the sports literature has demonstrated notable success in improving visual functions through training (Page et al., 2011; Schwab and Memmert, 2012; Rezaee et al., 2012; Krzepota et al., 2015). The theoretical basis for lifeguard-specific training follows principles of perceptual learning, where repeated exposure to domain-specific stimuli enhances the identification of critical features (Kellman and Garrigan, 2009). Through targeted practice, individuals develop automaticity in recognizing important patterns, reducing attentional demands and improving performance. In lifeguarding contexts, this involves learning to quickly identify the characteristic movements and postures that indicate drowning amidst numerous distractors. In the field of lifeguard drowning detection, a wide array of methodological approaches has been employed, including the analysis of authentic real-world recordings of drowning events, examination of pre-recorded footage, and the utilization of animated stimuli (Lanagan-Leitzel and Moore, 2010; Lanagan-Leitzel, 2012; Page et al., 2011; Laxton et al., 2021; Laxton et al., 2022; Laxton et al., 2021; Laxton & Crundall, 2018; Smith et al., 2020; Sharpe et al., 2023). While these methodological variations have provided valuable insights, they have also been identified as a key factor contributing to the conflicting findings in the existing literature (Smith, 2016). For example, the adopted tasks include varying numbers of visible patrons, fluctuating numbers of bathers, water-based obstacles, environmental stimuli, and task durations do not align with lifeguard responsibilities (Smith et al., 2020).

To address these limitations, a controlled animated task has been developed that replicates the core context of a lifeguard's role, encompassing the environment, duration, and task requirements, while excluding task-irrelevant variables (Sharpe et al., 2023, 2024). This animated task provides researchers with the opportunity to investigate various factors that contribute to successful drowning detection performance across a wide range of scenarios (i.e., bather count, drown durations), which are often impractical to replicate in a real-world setting. However, the use of animated tools or simulators is not without scrutiny, as highlighted by previous research (Fisher et al., 2011). Nevertheless, these tools possess inherent appeal due to their ability to incorporate human factors into research designs and offering advantages such as time efficiency, cost savings, and ethical considerations. As such, animated tasks provide researchers with the opportunity to investigate various factors that contribute to successful drowning detection performance. To date, these animated tools have shed light on the limitations of human detection performance, vigilance, and the disparities associated with lifeguard expertise (Page et al., 2011; Sharpe et al., 2023). Yet, the extent to which data generated from these animated tools can be generalized to replicate real-world environments remains uncertain. Exploring the validity of animated tasks through diverse methodologies, including assessments of external validity, plays a vital role in examining the degree to which laboratory tasks accurately represent real-world conditions or determine the generalizability of observed causal relationships (Blana and Golias, 2002; Branzi et al., 2017; Godley et al., 2002; McWilliams et al., 2018; Blaauw, 1982; Groeger and Murphy, 2020; Törnros, 1998; Kihlstrom, 2021).

1.4. Study aims and methods

The present research consists of two interconnected experiments designed to address critical gaps in the lifeguarding literature. First, we examine the external validity of animated drowning detection tasks compared to real-world scenarios. Second, we investigate whether training interventions—both domain-specific and domain-general—can enhance drowning detection performance in lifeguards.

Experiment 1 aims to assess the external validity of an animated lifeguard tool in comparison to a real-world counterpart matched in terms of duration, number of drownings, drowning times (i.e., the time of the drown events), and drowning durations. Specifically, we examine the impact of lifeguard experience and time on lifeguard-specific drowning detection performance while assessing the subjective and external validity of the animated tool. Building on previous literature, we hypothesize that lifeguards will outperform individuals with no lifeguarding experience, and that a decline in performance will be observed over time, as indicated by established findings on vigilance decrements (Risko et al., 2012; See et al., 1995; Thiffault and Bergeron, 2003; Verster and Roth, 2013). Based on prior literature utilizing the animated tool (Sharpe et al., 2023, 2024) and the similarity in outcomes to prior research in terms of expertise effects, performance declines, and cognitive load (Lanagan-Leitzel and Moore, 2010; Page et al., 2011; Laxton et al., 2021; Laxton et al., 2022; Laxton et al., 2021; Laxton & Crundall, 2018), we anticipate that the animated tool and real-world variant will show similar patterns of performance (i.e., absolute validity).

Experiment 2 aims to determine the influence of Lifeguard Specific Training (LST, utilizing the animated tool) and Working Memory (WM) training on WMC and drowning detection performance (utilizing the real-world recording). Previous research has consistently shown that experience positively correlates with drowning detection performance (Lanagan-Leitzel and Moore, 2010; Laxton et al., 2021; Vansteenkiste et al., 2021). Therefore, we hypothesize that regular exposure to simulated drowning scenarios will enhance real-world performance, leading to increased detection rates. Additionally, given that WMC has been shown to predict lifeguard-specific vigilance performance (Sharpe et al., 2024), we further hypothesize that drowning detection performance will improve following WM training. This prediction is based on the potential for WM training to enhance attentional control processes that are known to modulate individual differences in WMC (Engle and Kane, 2004; Kane et al., 2007; Kane and Engle, 2002). By linking these two experiments, we aim to provide a comprehensive understanding of both the validity of lifeguard training tools and the effectiveness of different training approaches for improving drowning detection abilities.

1.5. Experiment 1 methods

1.5.1. Participants

A total of 32 participants aged 18–38 years (M age = 25.19, SD = 4.99 years), consisting of 11 females and 21 males, took part in Experiment 1. Participants were divided into two groups: 16 experienced lifeguards (M lifeguard employment = 120.44, SD = 26.25 months) and 16 non-lifeguards with no experience. Following Sharpe et al. (2023), experienced lifeguards were defined as those with more than 100 months of lifeguarding experience. The lifeguard group had been personally involved in active rescues that would have otherwise led to full submersion (M active rescues = 0.94, SD = 1.48) and had provided poolside support during drowning incidents (M passive rescue = 4.69, SD = 5.68). The non-lifeguard group had no experience with drowning events. All lifeguards were actively employed as poolside lifeguards at the time of the study. Sample size was determined using G*Power 3.1.9.4 software (Faul et al., 2007) based on the experience by vigilance interaction effect size ($\eta p^2 = 0.235$) reported by Sharpe et al. (2023). For a power (1-β) of 0.95 and a two-tailed α of 0.05, the required sample size was n = 32. All participants were recruited through a United Kingdom lifeguard organisation via word of mouth. Ethical approval was obtained from the lead institutions ethical research committee [2212_13], and all participants provided informed consent.

1.5.2. Materials and tasks

Two drowning detection tasks were used: an animated task ("Bobbing Along") and a real-world recording. Both tasks were designed to be comparable in duration (30 min), number of drowning events (6), drowning locations, and drowning durations.

Lifeguard Specific Drowning Detection Tool. The lifeguard-specific drowning detection tool utilized in this study consisted of a segment derived from a previously adopted paradigm, titled *"Bobbing Along"* (see Sharpe et al., 2023 for additional detail). The tool presented 16 bathers within an aquatic environment for a duration of 30 min. Creation of the task was facilitated through Unreal Engine 4 (UE4), employing customized C++ code to establish the necessary functionality for a standard paradigm task. The development process also incorporated built-in blueprints, which streamlined the creation and monitoring of the 3D environment (Hill, 2021). For this study, one video was adopted, depicting a single scenario. The environment was divided into 16 navigation meshes, with one actor (i.e., bather) allocated to each mesh varying based on the specific task condition (refer to Fig. 1). The actors exhibited randomized movement patterns within their respective meshes, simulating swimming behaviour. In the event of a simulated drowning incident, a pre-determined bather would transition from treading water to gradually submerging over a specified period, adhering to the concept of the Instinctive Drowning Response (Pia, 1974). This passive drowning behaviour was selected to avoid the

Fig. 1. Screen capture of the "Bobbing Along" task as observed by the participant as seen on a projector.

inclusion of active drowning behaviours, such as flailing arms, splashing, or gasping for air, to prevent potential bias in lifeguards' attention towards specific taught behaviours that could influence their detection performance (Carballo-Fazanes et al., 2020). Throughout the 30-min task, there were no restarts, pauses, or alterations in the positions of the bathers. Once a bather had fully submerged, they would resurface after a 10-s interval and resume their randomized swim pattern. Consequently, there was no opportunity for delayed detection of drowning behaviour exceeding 10 ss. The swim patterns, drowning locations, and timings were consistent across all participants. The continuous nature of the task aimed to emulate the real-world responsibilities of lifeguards, who are tasked with monitoring all bathers concurrently within an aquatic environment.

Real-World Recording. The lifeguard-specific drown detection tool employed in this study simulated a real-world pool environment and included a series of bathers and drowning scenarios (refer to Fig. 2). The task was designed by scripting a team of 13 bathers and 3 certified lifeguards to emulate the behaviours described in the study "Bobbing Along" by Sharpe et al. (2023). The pool environment was divided into 16 navigation meshes, with the certified lifeguards assigned specific meshes where they would attempt to stay within while treading water or performing a relaxed breaststroke, replicating the drown locations depicted in the animated tool. The remaining bathers were free to swim using the breaststroke or tread water within their designated meshes. During the task, when a drowning event occurred, a pre-determined bather (a certified lifeguard) would begin treading water and position themselves face down for precisely 30 s. The task duration was 60 min, and there were no restarts, pauses, or changes in bather positions throughout this time. After the 30-s interval, the bather would resurface and continue their designated swim pattern. Unbeknownst to the participants, a bather would drown every 5 min during the 30-min duration of the task. It is important to note that the lifeguard-specific drown detection task implemented in this study followed a fixed timeline and did not involve any interruptions or modifications to the positions of the bathers.

NASA-Task Load Index. The measurement of participant perceived workload in this study was conducted using the NASA-Task Load Index (NASA-TLX), originally developed by Hart and Staveland (1988) and widely regarded as a robust measure of perceived mental workload (Rubio et al., 2004). The NASA-TLX assesses workload perception across six dimensions: mental demand (level of mental activity required), physical demand (level of physical activity required), temporal demand (sense of time pressure), performance (perceived level of success), effort (degree of exertion expended), and frustration (feelings of insecurity, discouragement, stress, or irritation). Participants were asked to rate each dimension on a scale of 0–100, with increments of 5. In line with the methodology employed by Sharpe et al. (2023), and for the benefit of comparison, we focused our analysis solely on the dimensions of

Fig. 2. Screen capture of the real-world recording as observed by the participant as seen on a projector.

effort, mental demand, and frustration. The internal consistency reliability (Cronbach's α) of the scales ranged from 0.87 to 0.93, indicating strong reliability in measuring participants' perceptions of workload in these dimensions.

1.5.3. Procedure

All testing sessions were conducted on consecutive days during regular working hours (7 a.m.–6 p.m.), and the tasks were presented in a random order to mitigate potential order effects. Participants were scheduled for testing at the same time for each session (e.g., Monday 1 p.m. and Tuesday 1 p.m.). To prevent vigilance decrement associated with multiple testing, participants engaged in only one task per day, and their second day of testing took place at the same designated time as the previous day. Prior to the commencement of the tasks, participants completed the necessary procedures, including the consent form and demographic questionnaire. Additionally, they were provided with a practice trial to ensure their comprehension of the target stimuli (i.e., drown event) and their ability to clearly perceive the display. The practice trials consisted of a 1-min segment of the task, during which a bather initiated the drowning event. All participants successfully detected the drown event during the practice trial without any prompting. Following the practice trial, participants had the opportunity to ask questions and then took their seat, maintaining 2 m from the projector screen. Participants were instructed to indicate if they observed a drowning event unfolding by using a response clicker, which allowed the researcher to record Hits (correctly detecting the drown scenario) or False Alarms (responding to a stimulus that was not present). Failure to respond to the drowning event was recorded as a Miss. Participants had the freedom to provide multiple responses and verbally express their observations (e.g., "a drown is occurring in the bottom left of the scene"). A researcher was present throughout the testing to ensure the accuracy of these detections, thereby minimizing the possibility of responding to false alarms during actual drowning events. As each task consisted of six drown events, the total number of successful Hits (where Hits were assigned a value of 1 and Misses a value of 0) enabled the researchers to calculate a Performance Score ranging from zero to six.

The tasks were presented to participants at 2 m on a high-definition (4K) SAMSUNG widescreen 16:9 projector, measuring 16 feet by 9 feet. The visual display was connected to a Reign gaming computer equipped with a GEFORCE GTX 1650 graphics card. Unbeknownst to the participants, drown events occurred at 5-min intervals throughout the tasks, with a total of six drown events pre-established at random locations. The drown locations were not arranged in a linear pattern but rather selected randomly (e.g., back middle, front left, middle right, etc.). Both the animated task and the real-world recording included drown events at similar locations and timings, ensuring consistency (e.g., the first drown event occurred at the bottom left of the scene in both tasks). Each participant viewed an identical version of each task, with no knowledge of the number of drown events embedded within the tasks. Apart from the presence of the researcher, participants completed the tasks individually in a quiet room illuminated by artificial lighting. The room remained darkened from natural light so that illumination could be controlled (M Horizonal $= 6.34$, $SD = 1.234$ Lx; M Vertical $= 15.75$, $SD = 2.55$ Lx) across all testing (recorded through the LUX LIGHT APP). Room temperature and humidity was controlled throughout all testing ($M_{\text{Temperature}} = 21.63^{\circ}$, $SD = 0.82$; $M_{\text{Humidity}} = 44$ %, $SD = 6.55$). On completion of each task, participants were asked to complete the NASA-TLX. The task length remained constant at precisely 30 min for each task with no interruptions (i.e., the participant continued to monitor the aquatic space for the entirety of the tasks).

1.5.4. Data analysis

Variables were assessed for univariate normality using skewness and kurtosis ratios (Fallowfield et al., 2005), with all measures falling within acceptable ranges (Kline, 1998). Boxplots revealed no univariate or multivariate outliers. A three-way repeated measures ANOVA examined the effects of group (Experienced vs. Non-lifeguard), task (Animated vs. Real-World), and time (6 drowning scenarios) on drowning detection performance (number of correct detections). Task differences in overall performance were assessed using paired samples t-tests. Two-way repeated measures ANOVAs analysed the effects of group and task on NASA-TLX workload subscales. Post-hoc repeated measures ANOVAs explored simple effects with Bonferroni adjustments to control for Type I errors (McLaughlin and Sainani, 2014). Violations of sphericity were addressed when epsilon was greater than 0.75 (Girden, 1992). Statistical significance was set at $\alpha = 0.05$, with effect sizes reported as partial eta squared (ηp^2) for ANOVA and Cohen's d for pairwise comparisons (Cohen, 1988).

1.6. Experiment 2 methods

1.6.1. Participants

A total of 48 novice lifeguards (M lifeguard employment = 30.06, SD = 16.50 months) aged 19–25 years (M age = 20.96, SD = 2.031 years), including 22 females and 26 males, participated in Experiment 2. Following Sharpe et al.'s (2023) criteria, all participants were categorized as novice lifeguards (less than 100 months of certified experience). Participants were randomly assigned to one of three groups: Lifeguard Specific Training (LST) group: 16 lifeguards (M lifeguard employment = 34.25, SD = 13.538 months; M age = 21.06, SD = 2.02 years); Cognitive Training (CT) group: 16 lifeguards (M lifeguard employment = 27.00, SD = 18.257 months; M age = 20.31, SD = 2.00 years); and Control group: 16 lifeguards (M lifeguard employment = 28.94, SD = 17.52 months; M age = 21.50, SD = 2.00 years). Group sizes were determined based on previous experimental literature showing large transfer effects associated with WM training (Nikravesh et al., 2021) and exceeded the average sample size reported in prior meta-analyses (Au et al., 2015). Initially, 54 lifeguards were recruited, but six were unable to continue beyond baseline assessment. All participants were actively employed as pool lifeguards during the study. Ethical approval for Experiment 2 was incorporated within the prior mentioned application [2212_13], and all participants provided informed consent.

1.6.2. Lifeguard specific assessment

The real-world recording task from Experiment 1 was used to assess drowning detection performance. The 30-min task featured six drowning events at 5-min intervals in pre-established random locations. Participants used a response clicker to indicate observed drowning events, with researchers recording Hits and Misses. A performance score was calculated based on the total number of successful Hits (range: 0–6).

1.6.3. Cognitive assessment

Three classic cognitive tasks were used to assess working memory capacity. These tasks collectively generated a composite WMC score as illustrated in Sharpe et al. (2024).

Spatial Span. Derived from the Corsi Block Tapping Task (Corsi, 1972), this task evaluated spatial short-term memory capacity and is a popular measure of WMC when combined with the below assessments (Metzler-Baddeley et al., 2017). Sixteen squares in a 4x4 grid were sequentially flashed (one every 900ms) in randomized order. Participants reproduced the sequence by clicking the squares in order of appearance. Difficulty increased from four flashes per round based on success rate. The test concluded after three errors.

Forward and Backward Digit Span tasks. Adapted from the verbal working memory component of the WAIS-R intelligence test (Weschler, 1981) both tasks were administered. Forward recall assesses the basic storage capacity of the phonological loop, while backward recall, which necessitates storage and manipulation of information prior to recall, engages visuospatial short-term working memory. These tasks have been widely employed, individually or collectively, in prior research to measure WMC (Lee et al., 2007; St Clair-Thompson and Gathercole, 2006). Like the Corsi Block Tapping task, participants observe a sequence of digits displayed sequentially on the screen. They are required to recall either the original linear sequence or its reversed order. Difficulty dynamically adjusts in each round, with a digit added or removed based on previous performance. The test concludes after three errors. All three memory span tasks contribute to the generation of a composite WMC score (as illustrated in Sharpe et al., 2024).

1.6.4. Lifeguard Specific Training (LST)

The LST group took part in one 30-min drowning detection training session per week for 4 weeks. For practical purposes, the LST group were split four smaller focus groups ($n = 4$ per group). Each session required access a study link sent through email. All LST was held remotely via Zoom video teleconferencing. To maintain participant anonymity, participant identities were hidden behind a participant ID (e.g., P1, P2, P3) and were not able to active their webcams. Training instructions and the task were demonstrated by the trainer through screensharing. Here, the trainer adopted the 'bobbing along' tool (Sharpe et al., 2023) to present a range of drown scenarios. Specifically, each session involved a series of simulated drown events with manipulated parameters (e.g., varying bather numbers, drown durations, and task lengths). Consisting of nine task variations, including 16, 32, and 68 bathers with 10, 30, and 90 s drown durations, participants observed a total of 96 drowning locations that varied in objective difficulty (as determined by Sharpe et al., 2023). Participants were shown segments from each variation and were asked to determine how many drownings occurred, at what locations, and at what times. All participants observed the same weekly material. Participants were required to note this information with pen and paper. Each segment lasted 5 min. If a drowning scenario was missed, then the footage was repeated to show the exact location of the drown. Locations were covered across all areas of the animated environment. With 24 drowning events being demonstrated and outlined per session, all training sessions lasted exactly 30-min. By 'outlined', the trainer merely stated the time and location of each drown. As no prior lifeguard literature has provided empirically justified instructions for the training of lifeguard drowning detection performance, the training sessions included no advice or direction regarding what to look for, how to look, or potential methods for remaining attentive.

1.6.5. Cognitive training (CT)

The CT group took part in one 30-min drowning detection training session per week for 4 weeks. Each session required participants to take part in a dual n-back WM training task (Jaeggi et al., 2008) through Millisecond Inquisit 6 Web application. Inquisit allowed participants to freely download the associated training package on a computer devise and location of their choosing. The task is a widely adopted research tool in cognitive training studies (Beloe and Derakshan, 2020; Owens et al., 2013; Sari et al., 2016; Swainston and Derakshan, 2018). The task presented a 3x3 grid, with a fixation point in the centre of the display. Per trial a green square stimulus displayed at a random location within the grid, in conjunction one of eight possible consonants (i.e., c, h, k, l, q, r, s, and t) was audibly presented within 500 ms. Participants were required to recall and correctly respond to the location of the stimulus and/or the letter that was heard depending on the type of the sequence (e.g., n = 1, 2, 3, or 4). Responses were recorded by participants pressing A-key for a visual match and S-key for an auditory match. In some instances, no response was required. Participants were instructed to respond as quickly as possible. The delay between trials was set at 2500 ms and the task could not be paused once starting. Each block lasted approximately 1 min, with a 15 s break between blocks. As the training task was adaptive (i.e., the n-back level was adjusted based on success), each participant performed varying levels of task difficulty. Irrespective of performance capability, all training sessions began with a single block at 1-back. Training included 20 blocks of 20 + n trials per session. As commonly adopted (Owens et al., 2013), the task's n-level increased by one n per block if accuracy for both modalities was equal or greater than 95 %, decreased if accuracy fell below 75 %, and remained if accuracy appeared stable (between 75 % and 95 %). As Jaeggi et al. (2008) demonstrated 4-back was the average level achieved across a typical training period, the highest level possible was set at 4-back.

1.6.6. Procedure

Upon entering the laboratory, participants were asked to read the information sheet, complete the consent form, and demographic questionnaire (i.e., age, gender, lifeguard experience). Following a single-blind procedure, participants were there randomly assigned into groups. Irrespective of the assigned groups (LST, CT, or control), all participants then underwent a lifeguard specific and cognitive assessment. All testing was carried out within normal working hours (8am - 4pm), with each task presented in a random order to avoid potential order effects. Other than the researcher, participants completed all assessment tasks alone, in a quiet, and artificially lit room. The room remained darkened from natural light so that illumination could be controlled (*M* Horizonal = 9.66, *SD* = 1.003 Lx; *M* Vertical = 14.29, *SD* = 3.10 Lx) across all testing (recorded through the LUX LIGHT APP). The assessment visits took approximately 55-min to complete. For LST and CT groups, training began on the day following the initial assessment. Post-assessments were held one day after a participants final training day. For all groups, testing was carried out over a period of 23 days. The data from participants that were unable to attend the laboratory for post-assessment testing were removed from the study. After post-assessment, participants were asked if they would be willing to be contacted for additional testing. This was asked for the purpose of a two-week retention test, if significant drowning detection performance differences at baseline to post-assessment were observed.

For the lifeguard specific assessment, participants first observed a practice trial of a scripted drown scenario to ensure they understood the target stimuli (i.e., drown event) and could clearly see the display. The practice trials consisted of a 1-min segment of task where a bather begins to drown from the onset of the video. All participants detected the drown event within the practice trial without prompting. Following the practice trial participants were given time to ask questions and be seated (2 m away from the projector screen). Performance was then recorded as per Experiment 1, with an identical environment and procedure followed. As prior, all drown events occurred at 5-min intervals in a pre-established location consisting of six drown events. Drown locations were selected at random (i.e., back middle, front left, middle right etc.) and did not follow a liner path (e.g., front, middle, and then back). Each participant observed an identical version of each task. Participants were unaware of the number of drown events occurring throughout the tasks. The task remained constant at precisely 30 min with no interruptions (i.e., the participant continued to monitor the aquatic space for the entirety of the tasks).

The cognitive assessment comprised of three tasks, based on classical paradigms from the cognitive neuroscience literature, to collectively record working memory capacity (as seen in Sharpe et al., 2024). Such tasks were designed and programmed by A.H. and have been utilized in previous studies (Corbett et al., 2015; Daws and Hampshire, 2017; Hampshire et al., 2012, 2019, 2021, 2022; Metzler-Baddeley et al., 2017; Owen et al., 2010). All cognitive testing was completed on Reign gaming computer (GEFORCE GTX 1650). The cognitive tasks took on average 16 min (*SD* = 2.65) to complete with no interruptions.

1.6.7. Data analysis

Data for each variable were screened for univariate normality (Fallowfield et al., 2005) and met normality criteria (Kline, 1998). Boxplots revealed no univariate or multivariate outliers. A one-way between-subjects ANOVA analysed group differences in WMC at baseline. A two-way repeated measures ANOVA examined the effect of group and assessment (Baseline vs. Post-training) on WMC. A paired-samples *t*-test analysed differences in the CT group's mean n-back level between first and final training sessions. For drowning detection performance, a two-way repeated measures ANOVA analysed group and time (6 drowning scenarios) differences at baseline. A three-way repeated measures ANOVA examined the effects of group, assessment, and time on drowning detection performance. A paired-samples *t*-test analysed differences in the LST group's mean performance between post-training and two-week retention assessments. Statistical significance was set at $\alpha = 0.05$, with effect sizes reported as partial eta squared (ηp^2) for ANOVA and Cohen's d for pairwise comparisons (Cohen, 1988).

2. Results

2.1. Experiment 1 results

2.1.1. Drowning detection performance

Main Effects. There was a significant main effect of experience group on total drowning detection performance ($F(1, 30) = 49.162$, $p < .001$, $\eta p^2 = .621$). The most experienced group ($M = 4.69$, $SD = 0.84$) performed on average greater than the naïve group ($M = 2.72$, $SD = 0.88$). There was a significant main effect of performance across time points ($F(4.227, 126.817) = 32.932$, $p < .001$, $\eta p^2 = .523$). On average performance began to deteriorate as time progressed. When averaging over levels of experience group and time, the animated task ($M = 3.94$, $SD = 1.32$) and real-world task ($M = 3.50$, $SD = 1.30$) demonstrated a statistically significant differences in total drowning detection performance ($t(31) = 3.304$, $p < .01$, $d = 0.584$).

Interaction Effects. Task had no 2-way interaction with time ($F(3.817, 114.495) = 0.901$, $p > .05$), nor 3-way interaction effects with time and experience group ($F(3.817, 114.495) = 1.570$, $p > .05$) on performance. Experience group had a significant interaction with time ($F(4.227, 126.817) = 6.948$, $p < .001$, $\eta p^2 = .188$; Fig. 3). Two separate post hoc repeated measure ANOVAs demonstrated the differences in performance scores across time points were significant for experienced ($F(5, 75) = 19.043$, $p < .001$, $\eta p^2 = .559$) and naïve groups ($F(5, 75) = 20.574$, $p < .001$, $\eta p^2 = .578$). Those with greater experience maintained their performance for longer periods of time.

2.1.2. Subjective workload

Main Effects. There was a significant main effect of experience group on subjective mental demand overall ($F(1, 30) = 20.034$, $p < .001$, $\eta p^2 = .400$). The experienced group ($M = 23.59$, $SD = 14.79$) reported the tasks as less mentally demanding than the naïve group ($M = 50.47$, $SD = 19.34$). There was a significant main effect of task on mental demand ($F(1, 30) = 32.932$, $p < .001$, $\eta p^2 = .605$). On average subjective mental demand was greater for the real-world task (M = 45.16, SD =

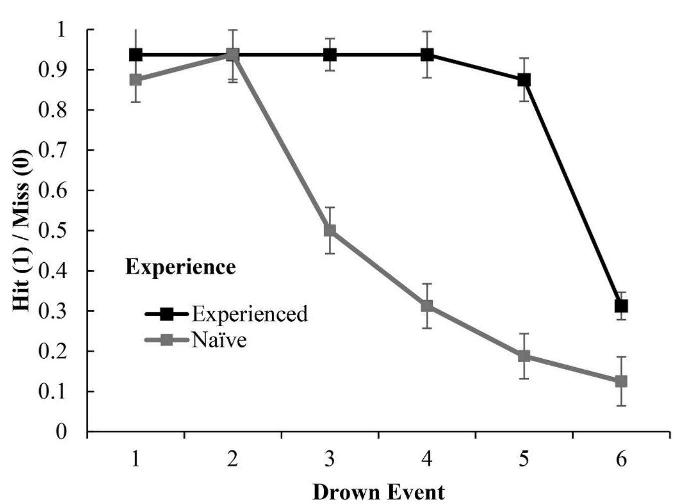

Fig. 3. The influence of experience and time on drown detection performance (with SE bars). Drown events occurred every 5 min (e.g., 1 = 5 min, 2 = 10 min …).

21.68), compared to the animated task (M = 28.91, SD = 23.92). There was a significant main effect of experience group on subjective frustration overall (*F(1, 30)* = 9.056, $p < .01$, $\eta p^2 = .232$). The experienced group (*M* = 17.19, *SD* = 4.553) reported the tasks as less frustrating than the naïve group (*M* = 31.86, *SD* = 15.65). There was a significant main effect of task on frustration (*F(1, 30)* = 24.875, $p < .001$, $\eta p^2 = .453$). On average subjective frustration was greater for the real-world task (*M* = 31.72, *SD* = 18.78), compared to the animated task (*M* = 17.34, *SD* = 6.06). There was a significant main effect of experience group on subjective effort overall (*F(1, 30)* = 16.642, $p < .001$, $\eta p^2 = .357$). The experienced group (*M* = 51.82, *SD* = 14.79) reported the tasks requiring less effort than the naïve group (*M* = 94.38, *SD* = 15.57). There was no significant main effect of task on effort (*F(1, 30)* = 2.074, $p > .05$).

Interaction Effects. Task had a 2-way interaction effect with experience group (*F(1, 30)* = 7.487, $p < .05$, $\eta p^2 = .200$) on mental demand. Subjective mental demand was reported higher for the real-world task, compared to the animated task, for lifeguards ($t(15) = 6.726$, $p = .001$) and the naïve group ($t(15) = 2.856$, $p = .01$). Task did not share a 2-way interaction effect with experience group on frustration (*F(1, 30)* = 0.423, $p > .05$) or effort (*F(1, 30)* = 1.389, $p > .05$).

2.2. Experiment 2 results

2.2.1. Cognitive assessment

All cognitive measures were subjected to factor reduction simultaneously by conducting a principal component analysis (PCA) given the commonality of function shared between all cognitive assessment measures (Hedden and Yoon, 2006; Miyake et al., 2000; Unsworth and Robison, 2020). Recommendations were followed given the relatively small sample size (de Winter et al., 2009; Preacher and MacCallum, 2002); however, it was anticipated that only a single component would be extracted from PCA (e.g., Sharpe et al., 2024). Numerous approaches were adopted given the variety of decision rules available, including parallel analysis (Horn, 1965), Kaiser criterion (Kaiser, 1960), and inspection of Cattell's scree plots (Cattell, 1952). Orthogonal Varimax rotation was then applied to the component matrix (Howard, 2016; Kaiser, 1958). One factor was extracted from the PCA with eigenvalues greater than one (see Fig. 4), accounting for 51 % of the total variance in the cognitive assessment tasks. The latent variable of WMC was then produced for participants by regressing individual task scores onto the rotated component matrix. Factor loading included Digit Span = .808, Reverse Digit Span = .740, and Spatial Span = .563.

Main Effect. There was no statistically significant main effect of group on baseline WMC ($F(2, 45) = 0.050$, $p > .05$, $\eta p^2 = .002$). No statistically significant main effect of the cognitive assessment (baseline and post cognitive assessment) was observed on WMC overall ($F(1, 45) = 0.008$, $p > .05$).

Interaction Effect. There were no 2-way interaction effects between cognitive assessment and group on WMC ($F(2, 45) = 0.423$, $p > .05$). However, performance differences at post-cognitive assessment appear to be descriptively greater for the CT group (Fig. 5) compared to baseline, but not for LST and control groups.

Exploratory Analysis. A series of two-way repeated measures ANOVAs were employed to explore the effect of group (LST vs. CT vs. control) and cognitive assessment (baseline vs. post cognitive assessment) on individual cognitive tasks. Irrespective of findings, it's important to note the methodological concerns associated with investigating individual items. A statistically significant main effect of the cognitive assessment was observed on digit span overall ($F(1, 45) = 6.410$, $p < .05$, $\eta p^2 = .125$). There was a 2-way interaction effect between cognitive assessment and group on digit span performance ($F(2, 45) = 3.423$, $p < .05$, $\eta p^2 = .132$; Fig. 6). The LST group and control group demonstrated no significant differences in digit span performance from baseline to post-cognitive assessment ($t(15) = 1.307$, $p > .05$ and $t(15) = 0.843$, $p > .05$, respectively). The CT group, however, did demonstrate significant differences in performance from baseline (M = 7.5, SD = 2.09) to post-cognitive assessment (M = 8.81, SD = 1.33; $t(15) = 3.542$, $p < .01$). No statistically significant main effect of the cognitive assessment was observed on reverse digit span ($F(1, 45) = 0.029$, $p > .05$) or spatial span overall ($F(1, 45) = 0.621$, $p > .05$). Likewise, there were no 2-way interaction effects between cognitive assessment and group on reverse digit span ($F(2, 45) = 0.465$, $p > .05$) or spatial span performance ($F(2, 45) = 1.230$, $p > .05$). The authors wish to note that this exploratory analyse must be approached with caution.

2.2.2. Working memory training

For those that took part in the CT group, there was a statistically significant difference in mean n-back level between the first (*M* = 2.44, *SD* = 1.03) and final (*M* = 3.19, *SD* = 0.66) training session ($t(15) = 4.392$, $p = < 0.001$).

2.2.3. Baseline drowning detection performance

Main Effect. There was no statistically significant main effect of group on total drowning detection performance at baseline ($F(2, 45) = 0.344$, $p > .05$). Averaging across all groups, time had a statistically significant main effect on total drowning detection performance at baseline ($F(5, 225) = 24.484$, $p < .001$, $\eta p^2 = .373$). On average

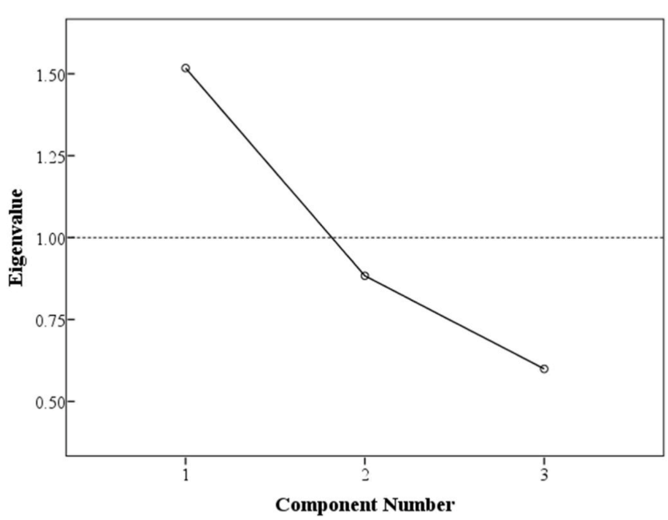

Fig. 4. Scree plot produced through PCA, demonstrating 1-component solution (eigenvalue >1), used for calculating individual latent cognitive variable. Dotted line highlights eigenvalue cut-off (Kaiser, 1960).

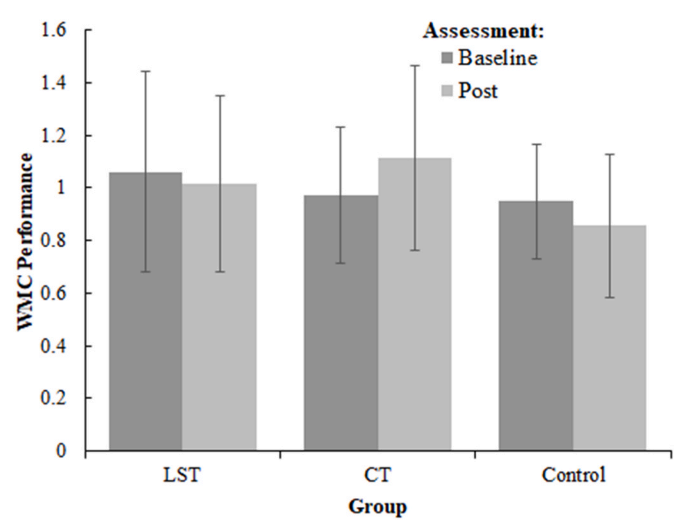

Fig. 5. The influence of group and cognitive assessment on mean WMC performance (with *SE* bars).

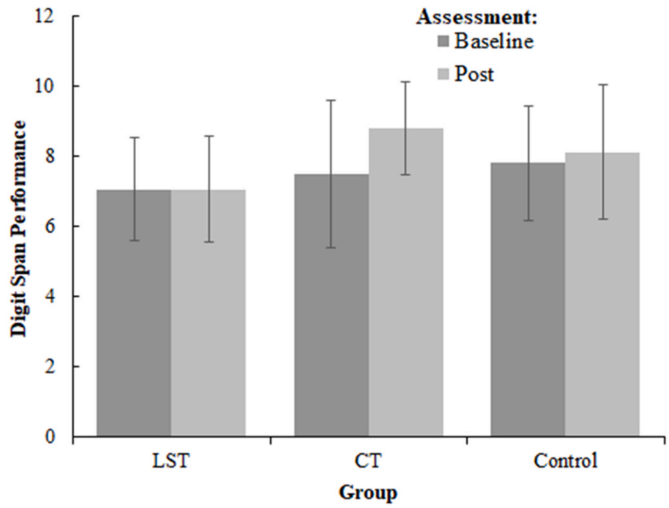

Fig. 6. The influence of group and cognitive assessment on mean digit span performance (with *SE* bars).

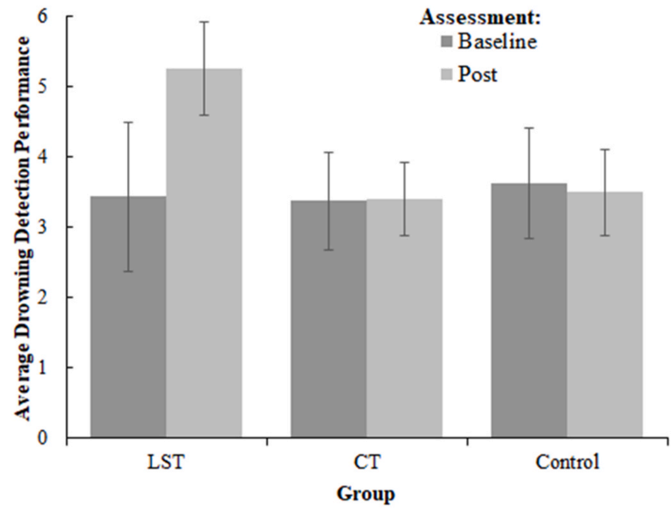

Fig. 7. The influence of group and lifeguard specific assessment on average drowning detection performance (with *SE* bars).

performance began to deteriorate as time progressed. From timepoint 1, participants experienced the first significant decline in performance at timepoint 4 ($t(47) = 4.057$, $p < .001$, $d = 0.627$).

Interaction Effect. There were no 2-way interaction effects between time and group on total drowning detection performance at baseline ($F(2, 45) = 0.344$, $p > .05$).

2.2.4. Lifeguard specific assessment

Main Effect. There was a statistically significant main effect of group on total drowning detection performance at post lifeguard specific assessment ($F(2, 45) = 13.402$, $p < .001$, $\eta p^2 = .373$). The LST group ($M = 5.25$, $SD = 0.68$) significantly outperformed the CT ($M = 3.188$, $SD = 0.54$) and control group ($M = 3.50$, $SD = 0.63$) at post assessment ($t(31) = 4.996$, $p < .001$, $d = 0.721$ and $t(31) = 3.674$, $p < .001$, $d = 0.530$ respectively). The CT and control group had no significant differences in post lifeguard specific assessment drowning detection performance ($t(31) = 1.323$, $p > .05$). A statistically significant main effect of the lifeguard specific assessment (baseline and post lifeguard specific assessment) was observed on total drowning detection performance overall ($F(1, 45) = 13.292$, $p < .001$, $\eta p^2 = .228$). Maintaining the trend, time had a statistically significant main effect on total drowning detection performance when averaging across lifeguard specific assessment and group ($F(5, 225) = 22.915$, $p < .001$, $\eta p^2 = .505$). Irrespective of any manipulated variable, performance consistently deteriorated as time progressed.

Interaction Effect. Lifeguard specific assessment had a 2-way significant interaction effect with group ($F(2, 45) = 22.915$, $p < .001$, $\eta p2 = .505$). The CT group and control group demonstrated no significant differences in performance from baseline to post lifeguard specific assessment ($t(15) = 0.789$, $p > .05$ and $t(15) = 0.526$, $p > .05$, respectively). The LST group was the only group to demonstrate significant differences in performance from baseline to post lifeguard specific assessment (M percentage difference = 29 %; $t(15) = 7.630$, $p < .001$). There were no 2-way interaction effects between time and group on drowning detection performance ($F(10, 225) = 0.889$, $p > .05$). There were no 2-way interaction effects between lifeguard specific assessment and time ($F(5, 225) = 0.142$, $p > .05$), and no 3-way interaction effect amongst variables on drowning detection performance ($F(10, 225) = 1.467$, $p > .05$). Whilst not sustained over time, LST held a significant positive influence on total drowning detection performance (Fig. 7).

2.2.5. Drowning detection retention

Only nine participants from the LST group were able to attend the two-week retention test. There was no statistically significant difference in mean drowning detection performance between the post-training assessment ($M = 5.44$, $SD = 0.72$) and two-week retention ($M = 5.00$, $SD = 0.47$; $t(8) = 1.315$, $p = > 0.05$).

3. Discussion

3.1. Experiment 1 discussion

The primary aim of Experiment 1 was to investigate the impact of lifeguard experience on drowning detection performance and assess the validity of an animated lifeguard tool by comparing it to a real-world recording. Consistent with our hypotheses and previous research (Lanagan-Leitzel and Moore, 2010; Laxton et al., 2021; Sharpe et al., 2023), experienced lifeguards demonstrated superior drowning detection performance compared to non-lifeguards. This suggests that lifeguard experience plays a crucial role in developing the ability to identify potential dangers in aquatic environments. Our findings indicate that experienced lifeguards exhibited better sustained attention over time across both tasks, aligning with previous literature suggesting that extended lifeguard experience may contribute to the development of attentional mechanisms that support vigilance (Sharpe et al., 2023). However, regardless of experience level, all participants experienced significant performance declines over time, consistent with findings from previous studies on vigilance decrements (Molley and Parasuraman, 2016; Risko et al., 2012; Swanson et al., 2012; Temple et al., 2000; Verster and Roth, 2013). These findings emphasize the limitations of human attention in extended monitoring tasks and tentatively support RLSS UK's recommendation for regular lifeguard rotations to mitigate vigilance decrements, though further research is needed.

Participants reported significantly higher levels of mental demand and frustration for the real-world task compared to the animated task, suggesting that the real-world task posed greater perceptual challenges. This may be attributed to greater visual complexity and less distinct figure-ground separation in the real-world recording. However, it is also possible that participants were more motivated to detect drowning events in the real-world task due to its naturalistic nature, as previous research has reported behavioural differences in response to real versus simulated stimuli (Ambadar et al., 2009). These differences suggest that the real-world task may have elicited greater attentional engagement but also increased cognitive load, possibly impacting performance. Further research should examine the specific visual and cognitive demands of real-world versus animated environments in lifeguard training. Contrary to expectations, experienced lifeguards did not report lower levels of perceived workload compared to non-lifeguards, with both groups reporting similar workload levels. This finding suggests that while experienced lifeguards perform better at drowning detection, they may still experience similar cognitive demands during the task. The alignment between perceived workload and task performance across both tasks indicates that as mental demand increased, performance decreased, highlighting the relationship between cognitive load and vigilance performance. Importantly, the animated tool demonstrated external validity, with no statistically significant difference in performance between the two tasks across lifeguard experience groups. The mean percentage performance differences between tasks were minimal (8 % for lifeguards and 7 % for non-lifeguards), and similar vigilance decrements were observed over time in both tasks. These results suggest that the animated tool effectively captures objective drowning detection performance in a manner comparable to its real-world counterpart, supporting its validity for training and assessment purposes. Future studies should explore whether further refinements, such as incorporating environmental distractors and real-world auditory elements, may enhance its applicability for training lifeguards in more complex and realistic settings.

3.2. *Experiment 2 discussion*

Experiment 2 aimed to determine the influence of Lifeguard Specific Training (LST) and Working Memory (WM) training on drowning detection performance and WMC. The cognitive training intervention showed no significant effects on either drowning detection performance or WMC, aligning with recent scepticism about WM training transfer effects (Moreau, 2021; Moreau et al., 2019; Simons et al., 2016) and contradicting earlier enthusiasm (e.g., Green et al., 2019). The absence of near-transfer effects to WMC suggests that the adaptive n-back task may not engage the same cognitive processes active during successful lifeguard drowning detection, despite its previously demonstrated improvements in inhibitory functions (Owens et al., 2013). This highlights the challenge of identifying which specific cognitive processes are critical for drowning detection and designing training protocols accordingly. Future research may benefit from recording differences in brain structure and function between expert and novice drowning detectors to better understand these processes. Alternatively, WM training might be more effective if it targets multiple cognitive processes rather than focusing on a single task (Kramer and Morrow, 2012). A multi-process approach could reduce the monotony associated with single-task performance and potentially increase the likelihood of positive training effects. As suggested by Gobet and Sala (2022), the field of cognitive training may need to temper expectations regarding far-transfer effects. The lack of impact from WM training raises important questions about which cognitive mechanisms are most crucial for improving lifeguard vigilance and whether alternative cognitive training approaches could be more effective. Future studies should explore task-specific cognitive interventions that may offer more direct improvements to lifeguard vigilance.

In contrast, the LST intervention, which exposed novice lifeguards to various drowning scenarios without explicit instruction, significantly improved drowning detection performance. This suggests that exposure to diverse drowning scenarios, even without detailed guidance, can enhance detection ability. The improvement may be attributed to increased familiarity with critical cues and the development of more efficient visual search strategies through repeated exposure to drowning events under varying conditions. These findings align with research indicating that repeated exposure to heterogeneous and unpredictable stimuli improves visual search accuracy (Schutster et al., 2013). They also support previous work emphasizing the role of familiarity with target features in improving performance in identification tasks (Fincannon et al., 2013). It is important to note, however, that the LST was closely aligned with the assessment task, potentially raising questions about whether performance improvements would transfer to different contexts or real-world scenarios. While our findings suggest that LST improves pool-based drowning detection, further research should explore whether similar benefits extend to open-water environments, where additional variables (e.g., waves, sun glare, dynamic patron movement) may complicate detection efforts. Furthermore, long-term retention of LST effects remains unclear, as some participants were unable to attend follow-up assessments. Given evidence from broader training literature suggesting that skill retention may decline over time without reinforcement (Miles et al., 2015), future studies should examine the durability of LST effects beyond a two-week retention period.

Additionally, the current study did not incorporate real-world distractors, such as poolside activity or auditory distractions, which are commonly present in lifeguard surveillance tasks. These contextual factors could influence real-world drowning detection performance and should be considered in future LST designs. Moreover, examining the interplay between LST and traditional lifeguard training methodologies may yield insights into optimizing surveillance training programs. For example, incorporating knowledge of results training (e.g., feedback on missed drowning events) or cognitive reinforcement techniques may enhance the long-term benefits of LST. Irrespective, Experiment 2 provides the first empirical evidence that LST can improve drowning detection performance in a controlled environment, reinforcing the value of scenario-based exposure in lifeguard training. While limitations exist regarding transferability to diverse aquatic settings, these findings underscore the need for training programs to prioritize surveillance-focused education. Future research should investigate the generalizability of LST to natural aquatic settings and determine whether complementary instructional methods can further enhance lifeguard vigilance and response efficiency.

3.3. *Real-world implications of findings*

The findings from this study have direct implications for lifeguard training, surveillance strategies, and drowning prevention efforts. The superior drowning detection performance of experienced lifeguards reinforces the necessity of structured, scenario-based training programs that expose trainees to varied drowning events. Empirical research tentatively supports this, with literature demonstrating that targeted training improves situational awareness and decision-making (Heusler and Sutter, 2022; Walshe et al., 2019). The observed decline in performance over time further underscores the importance of rotational monitoring strategies to counter vigilance decrements, consistent with workplace safety research showing that task-switching reduces cognitive overload and enhances hazard detection (Ren et al., 2023). The external validity of the animated training tool suggests that simulated environments can be effectively integrated into lifeguard training, paralleling research where simulation-based instruction enhances real-world performance (Salas et al., 2012). Additionally, the success of Lifeguard-Specific Training (LST) without explicit instruction highlights the value of experiential learning, aligning with studies showing that exposure to diverse scenarios improves recognition of critical cues (Benishek et al., 2015). However, the lack of transfer effects from working memory (WM) training indicates that cognitive interventions should be tailored specifically to lifeguard vigilance tasks, echoing broader findings that domain-specific training yields better performance improvements than generalized cognitive exercises (Harris et al., 2018). Future research should explore the integration of real-time feedback tools to enhance scanning behaviours and drowning detection efficiency (see Kirby, 2009 for an example in skiing). These findings reinforce the urgent need for comprehensive, evidence-based training methodologies to optimize lifeguard performance in real-world settings.

3.4. *Limitations and future research*

Several limitations should be considered when interpreting these findings. First, the study focused primarily on controlled, simulated environments, which may not fully capture the complexity and variability of real-world lifeguarding scenarios. Second, the participant samples lacked cultural and demographic diversity, potentially limiting the generalizability of the findings. Third, the relatively short duration of the training interventions (four weeks) may not have been sufficient to observe maximum training effects, particularly for cognitive training. Future research should explore the long-term retention of drowning detection skills, especially for novice lifeguards, and investigate how training effects transfer to different aquatic environments beyond swimming pools, such as open water or waterparks. Additionally, exploring multimodal training approaches that combine cognitive and perceptual-motor training might yield more comprehensive improvements in lifeguard performance. Alternative technologies such as virtual and augmented reality could provide more immersive and ecologically valid training environments. These technologies might offer advantages over traditional training methods by allowing for greater control over training parameters while maintaining high levels of perceptual fidelity. Finally, more research is needed to understand the specific cognitive and perceptual mechanisms underlying expert drowning detection performance, which could inform more targeted and effective training protocols. Likewise, authors must not ignore the possibility that it is possible that lifeguards, due to their training and experience, were more familiar with the critical cues associated with drowning. This familiarity may have allowed participants within this manuscript to identify drowning victims more efficiently than non-lifeguards, rather than solely relying on sustained attention. Perhaps it is plausible that the non-lifeguards lacked the necessary repertoire of visual cues required to detect drowning within the 30-s timeframe. Future research should investigate whether explicit training in drowning cues could enhance performance in naïve participants. Neuroimaging and eye-tracking methodologies could of course provide valuable insights into these mechanisms and how they develop with experience and training.

3.5. Recommendations for training instructions

Given that prior literature has demonstrated that lifeguards, irrespective of experience, have an inability to detect all drowning events across a lifeguard specific vigilance task (Sharpe et al., 2023), future intervention studies may wish to incorporate empirically justified guidance on individual sustained attention. Perhaps lifeguard researchers can look to cognitive literature to understand further the processes that contribute to the redirection of attentional focus (Kane et al., 2016; Unsworth & McMillan, 2013). Likewise, it appears lifeguards with practical and theoretical-based training, advanced water safety certification, but limited certificated lifeguard experience, have no detection advantage over those with no exposure to lifeguard employment (Lanagan-Leitzel and Moore, 2010). As our findings appeared to demonstrate exposure to drowning scenarios appears to elicit some benefit to drowning detection performance, authors could introduce a greater range of drowning scenarios of varying contextual differences (e.g., a range of bather numbers), environments (e.g., pool, beach, quarry, leisure), or drowning behaviours (e.g., active vs. passive). Finally, given the overwhelming support that the most experienced lifeguards consistently outperform those with lesser experience (Lanagan-Leitzel and Moore, 2010; Laxton et al., 2021; Page et al., 2011), and performance appears to decline rapidly as time progresses (Sharpe et al., 2023; Sharpe and Smith, 2024), perhaps the addition of empirically supported educational material may enable lifeguards to be more mindful of the limitations of human attention. Such knowledge may even lead to lifeguards requesting additional lifeguard support or breaks in-situ when they perceive their attention (i.e., ability to sustain attention) to be declining.

4. Conclusion

This study had two primary aims: (1) to evaluate the external validity of a previously used animated performance tool, focusing on the impact of lifeguard experience and task duration on performance measures, and (2) to investigate the effect of two distinct training protocols on lifeguard-specific drowning detection performance. Findings underscore the importance of incorporating drowning detection challenges, such as varying bather numbers, drowning durations, and locations, into lifeguard certification programs—an area currently underemphasized. Given that many lifeguards failed to detect all drowning scenarios at baseline, this study highlights the critical need for comprehensive drowning detection training. Future research should consider the utility of animated tools and focus on deepening our understanding of the mechanisms that support effective and sustained drowning detection.

CRediT authorship contribution statement

Benjamin T. Sharpe: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marcus S. Smith:** Writing – review & editing, Supervision. **Steven C.R. Williams:** Writing – review & editing, Supervision, Software. **Adam Hampshire:** Writing – review & editing, Software. **Maria Balaet:** Software. **William Trender:** Software. **Peter J. Hellyer:** Software. **Jo Talbot:** Resources. **Jenny Smith:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Au, J., Sheehan, E., Tsai, N., Duncan, G.J., Buschkuehl, M., Jaeggi, S.M., 2015. Improving fluid intelligence with training on working memory: a meta-analysis. Psychonomic Bull. Rev. 22 (2), 366–377. https://doi.org/10.3758/s13423-014-0699-x.
Baddeley, A., 2007. Working Memory, Thought and Action. Oxford University Press.
Bao, H., Fang, W., Guo, B., Qiu, H., 2024. Adaptive human–computer interface design for supervision task based on user attention and system state. Int. J. Hum. Comput. Interact. 40 (8), 2054–2066. https://doi.org/10.1080/10447318.2023.2228070.
Ball, K., Berch, D.B., Helmers, K.F., Jobe, J.B., Leveck, M.D., Marsiske, M., Morris, J.N., Rebok, G.W., Smith, D.M., Tennstedt, S.L., Unverzagt, F.W., Willis, S.L., for the ACTIVE Study Group, 2002. Effects of cognitive training interventions with older adults. JAMA 288 (18), 2271. https://doi.org/10.1001/jama.288.18.2271.
Ball, K., Edwards, J.D., Ross, L.A., 2007. The impact of speed of processing training on cognitive and everyday functions. J. Gerontol.: Ser. Bibliogr. 62 (Special_Issue_1), 19–31. https://doi.org/10.1093/geronb/62.special_issue_1.19.
Basak, C., Boot, W.R., Voss, M.W., Kramer, A.F., 2008. Can training in a real-time strategy video game attenuate cognitive decline in older adults? Psychol. Aging 23 (4), 765–777. https://doi.org/10.1037/a0013494.
Beloe, P., Derakshan, N., 2020. Adaptive working memory training can reduce anxiety and depression vulnerability in adolescents. Dev. Sci. 23 (4). https://doi.org/10.1111/desc.12831.
Brehmer, Y., Westerberg, H., Bäckman, L., 2012. Working-memory training in younger and older adults: training gains, transfer, and maintenance. Front. Hum. Neurosci. 6. https://doi.org/10.3389/fnhum.2012.00063.
Benishek, L.E., Weaver, S.J., Newman-Toker, D.E., 2015. The cognitive psychology of diagnostic errors. Scientific American Neurology 10 (7900.6288).
Buehner, M., Krumm, S., Ziegler, M., Pluecken, T., 2006. Cognitive abilities and their interplay. J. Indiv. Differ. 27 (2), 57–72. https://doi.org/10.1027/1614-0001.27.2.57.
Cattell, R.B., 1952. The three basic factor-analytic research designs—their interrelations and derivatives. Psychol. Bull. 49 (5). https://doi.org/10.1037/h0054245.
Chen, H., Liu, C., Zhou, F., Cao, X.Y., Wu, K., Chen, Y.L., et al., 2022. Focused-attention meditation improves flow, communication skills, and safety attitudes of surgeons. Int. J. Environ. Res. Publ. Health 19 (9), 5292. https://doi.org/10.3390/ijerph19095292.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Erlbaum.
Corbett, A., Owen, A., Hampshire, A., Grahn, J., Stenton, R., Dajani, S., Burns, A., Howard, R., Williams, N., Williams, G., Ballard, C., 2015. The effect of an online

cognitive training package in healthy older adults: an online randomized controlled trial. J. Am. Med. Dir. Assoc. 16 (11). https://doi.org/10.1016/j.jamda.2015.06.014.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3 (3), 201–215. https://doi.org/10.1038/nrn755.

Corsi, P.M., 1972. Human Memory and the Medial Temporal Region of the Brain. Unpublished Doctoral Dissertation. McGill University, Montreal.

Dahlin, E., Bäckman, L., Neely, A.S., Nyberg, L., 2009. Training of the executive component of working memory: subcortical areas mediate transfer effects. Restor. Neurol. Neurosci. 27 (5), 405–419. https://doi.org/10.3233/RNN-2009-0492.

Daws, R.E., Hampshire, A., 2017. The negative relationship between reasoning and religiosity is underpinned by a bias for intuitive responses specifically when intuition and logic are in conflict. Front. Psychol. 8. https://doi.org/10.3389/fpsyg.2017.02191.

Dehnabaei, Z., Tabibi, Z., Ouimet, M.C., Moghaddam, A.M., Delavar, M.E., 2024. Computerized cognitive training to improve executive functions and driving skills of adolescents with and without symptoms of attention-deficit/hyperactivity disorder. Transport. Res. F Traffic Psychol. Behav. 105, 13–23.

de Winter, J.C.F., Dodou*, D., Wieringa, P.A., 2009. Exploratory factor analysis with small sample sizes. Multivariate Behav. Res. 44 (2), 147–181. https://doi.org/10.1080/00273170902794206.

Edwards, J.D., Myers, C., Ross, L.A., Roenker, D.L., Cissell, G.M., McLaughlin, A.M., Ball, K.K., 2009. The longitudinal impact of cognitive speed of processing training on driving mobility. Gerontol. 49 (4), 485–494. https://doi.org/10.1093/geront/gnp042.

Engle, R., Kane, M.J., 2004. Executive attention, working memory capacity, and a two-factor theory of cognitive control. In: Ross, B.H. (Ed.), The Psychology of Learning and Motivation: Advances in Research and Theory, vol. 44. Elsevier Science, pp. 145–199.

Fallowfield, J.L., Hale, B.J., Wilkinson, D.M., 2005. Using Statistics in Sport and Exercise Science Research. Lotus Publishing.

Fernandez-Ballesteros, R., Botella, Zamarron, Cabras, Molina, Schettini, & Tarraga, 2012. Cognitive plasticity in normal and pathological aging. Clin. Interv. Aging 15. https://doi.org/10.2147/CIA.S27008.

Foroughi, C.K., Devlin, S., Pak, R., Brown, N.L., Sibley, C., Coyne, J.T., 2023. Near-perfect automation: investigating performance, trust, and visual attention allocation. Hum. Factors 65 (4), 546–561. https://doi.org/10.1177/00187208211032889.

Girden, E.R., 1992. ANOVA: Repeated Measures, Issue 84. Sage.

Gorman, A.D., Abernethy, B., Farrow, D., 2018. Reduced attentional focus and the influence on expert anticipatory perception. Atten. Percept. Psychophys. 80, 166–176. https://doi.org/10.3758/s13414-017-1429-z.

Greenlee, E.T., Hess, L.J., Simpson, B.D., Finomore, V.S., 2024. Vigilance to Spatialized Auditory displays: initial Assessment of performance and workload. Hum. Factors 66 (4), 987–1003. https://doi.org/10.1177/00187208221139744.

Hampshire, A., Chatfield, D.A., Mphil, A.M., Jolly, A., Trender, W., Hellyer, P.J., Giovane, M. Del, Newcombe, V.F.J., Outtrim, J.G., Warne, B., Bhatti, J., Pointon, L., Elmer, A., Sithole, N., Bradley, J., Kingston, N., Sawcer, S.J., Bullmore, E.T., Rowe, J. B., Menon, D.K., 2022. Multivariate profile and acute-phase correlates of cognitive deficits in a COVID-19 hospitalised cohort. EClinicalMedicine 47, 101417. https://doi.org/10.1016/j.eclinm.2022.101417.

Hampshire, A., Highfield, R.R., Parkin, B.L., Owen, A.M., 2012. Fractionating human intelligence. Neuron 76 (6). https://doi.org/10.1016/j.neuron.2012.06.022.

Hampshire, A., Sandrone, S., Hellyer, P.J., 2019. A large-scale, cross-sectional investigation into the efficacy of brain training. Front. Hum. Neurosci. 13. https://doi.org/10.3389/fnhum.2019.00221.

Hampshire, A., Trender, W., Chamberlain, S.R., Jolly, A.E., Grant, J.E., Patrick, F., Mazibuko, N., Williams, S.C., Barnby, J.M., Hellyer, P., Mehta, M.A., 2021. Cognitive deficits in people who have recovered from COVID-19. EClinicalMedicine 39, 101044. https://doi.org/10.1016/j.eclinm.2021.101044.

Harris, D.J., Wilson, M.R., Vine, S.J., 2018. A systematic review of commercial cognitive training devices: implications for use in sport. Front. Psychol. 9, 709.

Hedden, T., Yoon, C., 2006. Individual differences in executive processing predict susceptibility to interference in verbal working memory. Neuropsychology 20 (5). https://doi.org/10.1037/0894-4105.20.5.511.

Heusler, B., Sutter, C., 2022. Shoot or don't shoot? Tactical gaze control and visual attention training improves police cadets' decision-making performance in live-fire scenarios. Front. Psychol. 13, 798766.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30 (2). https://doi.org/10.1007/BF02289447.

Howard, M.C., 2016. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? Int. J. Hum. Comput. Interact. 32 (1). https://doi.org/10.1080/10447318.2015.1087664.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Perrig, W.J., 2008. Improving fluid intelligence with training on working memory. Proc. Natl. Acad. Sci. 105 (19), 6829–6833. https://doi.org/10.1073/pnas.0801268105.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Shah, P., 2011. Short- and long-term benefits of cognitive training. Proc. Natl. Acad. Sci. 108 (25), 10081–10086. https://doi.org/10.1073/pnas.1103228108.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Shah, P., 2012. Cogmed and working memory training—current challenges and the search for underlying mechanisms. Journal of Applied Research in Memory and Cognition 1 (3), 211–213. https://doi.org/10.1016/j.jarmac.2012.07.002.

Kaiser, H.F., 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika 23 (3). https://doi.org/10.1007/BF02289233.

Kaiser, H.F., 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20 (1). https://doi.org/10.1177/001316446002000116.

Kane, M.J., Conway, A., Hambrick, D., Engle, R., 2007. Variation in working memory capacity as variation in executive attention and control. In: Conway, A., Jarrold, C., Kane, M., Miyake, A., Towse, J. (Eds.), Variation in Working Memory. Oxford University Press, pp. 21–46.

Kane, M.J., Engle, R., 2002. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. Psychonomic Bull. Rev. 9 (4). https://doi.org/10.3758/BF03196323.

Karbach, J., Kray, J., 2009. How useful is executive control training? Age differences in near and far transfer of task-switching training. Dev. Sci. 12 (6), 978–990. https://doi.org/10.1111/j.1467-7687.2009.00846.x.

Kirby, R., 2009. Development of a real-time performance measurement and feedback system for alpine skiers. Sports Technol. 2 (1–2), 43–52.

Kline, R.B., 1998. Principles and Practice of Structural Equation Modeling. Guilford Press.

Lanagan-Leitzel, L.K., 2021. Does incident severity influence surveillance by lifeguards in aquatic scenes? Appl. Cogn. Psychol. 35 (1), 181–191. https://doi.org/10.1002/acp.3752.

Lanagan-Leitzel, L.K., Moore, C.M., 2010. Do lifeguards monitor the events they should? Int. J. Aquat. Res. Educ. 4 (3). https://doi.org/10.25035/ijare.04.03.04.

Laxton, V., Crundall, D., Guest, D., Howard, C.J., 2021. Visual search for drowning swimmers: investigating the impact of lifeguarding experience. Appl. Cogn. Psychol. 35 (1). https://doi.org/10.1002/acp.3756.

Laxton, V., Howard, C.J., Guest, D., Crundall, D., 2023. Intense classification training to increase the detection of drowning swimmers. Appl. Cogn. Psychol. 37 (1), 189–199.

Laxton, V., Mackenzie, A.K., Crundall, D., 2022. An exploration into the contributing cognitive skills of lifeguard visual search. Appl. Cogn. Psychol. 36 (1), 216–227. https://doi.org/10.1002/acp.3913.

Lee, Y., Lu, M., Ko, H., 2007. Effects of skill training on working memory capacity. Learn. InStruct. 17 (3). https://doi.org/10.1016/j.learninstruc.2007.02.010.

McLaughlin, M.J., Sainani, K.L., 2014. Bonferroni, Holm, and Hochberg corrections: fun names, serious changes to p values. PM & R : The Journal of Injury, Function, and Rehabilitation 6 (6), 544–546. https://doi.org/10.1016/J.PMRJ.2014.04.006.

Metzler-Baddeley, C., Foley, S., de Santis, S., Charron, C., Hampshire, A., Caeyenberghs, K., Jones, D.K., 2017. Dynamics of white matter plasticity underlying working memory training: multimodal evidence from diffusion MRI and relaxometry. J. Cognit. Neurosci. 29 (9). https://doi.org/10.1162/jocn_a_01127.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D., 2000. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. Cogn. Psychol. 41 (1). https://doi.org/10.1006/cogp.1999.0734.

Miyake, A., Shah, P., 1999. Models of Working Memory: Mechanisms of Active Maintenance and Executive Control. Cambridge University Press.

Nikravesh, M., Aghajanzadeh, M., Maroufizadeh, S., Saffarian, A., Jafari, Z., 2021. Working memory training in post-stroke aphasia: near and far transfer effects. J. Commun. Disord. 89, 106077. https://doi.org/10.1016/j.jcomdis.2020.106077.

Norman, D.A., Shallice, T., 1986. Attention to action. In: Consciousness and Self-Regulation. Springer US, pp. 1–18. https://doi.org/10.1007/978-1-4757-0629-1_1.

Owen, A.M., Hampshire, A., Grahn, J.A., Stenton, R., Dajani, S., Burns, A.S., Howard, R. J., Ballard, C.G., 2010. Putting brain training to the test. Nature 465 (7299). https://doi.org/10.1038/nature09042.

Owens, M., Koster, E.H.W., Derakshan, N., 2013. Improving attention control in dysphoria through cognitive training: transfer effects on working memory capacity and filtering efficiency. Psychophysiology 50 (3), 297–307. https://doi.org/10.1111/psyp.12010.

Pak, R., McLaughlin, A.C., Engle, R., 2024. The relevance of attention control, not working memory, in human factors. Hum. Factors 66 (5), 1321–1332. https://doi.org/10.1177/00187208231159727.

Page, J., Bates, V., Long, G., Dawes, P., Tipton, M., 2011. Beach lifeguards: visual search patterns, detection rates and the influence of experience. Ophthalmic Physiol. Opt. 31 (3). https://doi.org/10.1111/j.1475-1313.2011.00824.x.

Petersen, S.E., Posner, M.I., 2012. The attention system of the human brain: 20 Years after. Annu. Rev. Neurosci. 35 (1), 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525.

Preacher, K.J., MacCallum, R.C., 2002. Exploratory factor analysis in behavior genetics research: factor recovery with small sample sizes. Behav. Genet. 32 (2), 153–161. https://doi.org/10.1023/A:1015210025234.

Redick, T.S., Wiemers, E.A., Engle, R.W., 2020. The role of proactive interference in working memory training and transfer. Psychol. Res. 84 (6), 1635–1654. https://doi.org/10.1007/s00426-019-01172-8.

Ren, B., Zhou, Q., Chen, J., 2023. Assessing cognitive workloads of assembly workers during multi-task switching. Sci. Rep. 13 (1), 16356.

Rensink, R.A., 2002. Change detection. Annu. Rev. Psychol. 53 (1), 245–277. https://doi.org/10.1146/annurev.psych.53.100901.135125.

Robertson, I.H., Garavan, H., 2010. Vigilant attention. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences, first ed. MIT Press, pp. 631–640.

Robertson, I.H., Manly, T.O.M., Beschin, N., Daini, R., Haeske-Dewick, H., Hömberg, V., et al., 1997. Auditory sustained attention is a marker of unilateral spatial neglect. Neuropsychologia 35 (12), 1527–1532. https://doi.org/10.1016/S0028-3932(97)00084-5.

Salas, E., Tannenbaum, S.I., Kraiger, K., Smith-Jentsch, K.A., 2012. The science of training and development in organizations: what matters in practice. Psychol. Sci. Publ. Interest 13 (2), 74–101.

Sari, B.A., Koster, E.H.W., Pourtois, G., Derakshan, N., 2016. Training working memory to improve attentional control in anxiety: a proof-of-principle study using behavioral and electrophysiological measures. Biol. Psychol. 121, 203–212. https://doi.org/10.1016/j.biopsycho.2015.09.008.

Schweizer, K., Moosebrugger, H., 2004. Attention and working memory as predictors of intelligence. Intelligence 32 (4), 329–347. https://doi.org/10.1016/j.intell.2004.06.006.

Sharpe, B.T., Smith, J., 2024. Influence of Vigilance Performance on Lifeguard Gaze Behaviour. Europe's Journal of Psychology.

Sharpe, B.T., Smith, M.S., Williams, S.C.R., Hampshire, A., Balaet, M., Trender, W., Hellyer, P.J., Talbot, J., Smith, J., 2024. Cognition and lifeguard detection performance. Appl. Cogn. Psychol. https://doi.org/10.1002/acp.4139.

Sharpe, B.T., Smith, M.S., Williams, S.C.R., Talbot, J., Runswick, O.R., Smith, J., 2023. An expert-novice comparison of lifeguard specific vigilance performance. J. Saf. Res, *87*, 416-430.

Shipstead, Z., Hicks, K.L., Engle, R.W., 2012. Cogmed working memory training: does the evidence support the claims? Journal of Applied Research in Memory and Cognition 1 (3), 185–193. https://doi.org/10.1016/j.jarmac.2012.06.003.

St Clair-Thompson, H.L., Gathercole, S.E., 2006. Executive functions and achievements in school: shifting, updating, inhibition, and working memory. Q. J. Exp. Psychol. 59 (4), 745–759. https://doi.org/10.1080/17470210500162854.

Swainston, J., Derakshan, N., 2018. Training cognitive control to reduce emotional vulnerability in breast cancer. Psychooncology 27 (7), 1780–1786. https://doi.org/10.1002/pon.4727.

Unsworth, N., Robison, M.K., 2017. A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. Psychonomic Bull. Rev. 24 (4), 1282–1311. https://doi.org/10.3758/s13423-016-1220-5.

Unsworth, N., Robison, M.K., 2020. Working memory capacity and sustained attention: a cognitive-energetic perspective. J. Exp. Psychol. Learn. Mem. Cognit. 46 (1). https://doi.org/10.1037/xlm0000712.

Vansteenkiste, P., Lenoir, M., Bourgois, J.G., 2021. Gaze behaviour of experienced and novice beach lifeguards – an exploratory in situ study. Appl. Cogn. Psychol. 35 (1). https://doi.org/10.1002/acp.3760.

Walshe, N.C., Crowley, C.M., O'Brien, S., Browne, J.P., Hegarty, J.M., 2019. Educational interventions to enhance situation awareness: a systematic review and meta-analysis. Simulat. Healthc. J. Soc. Med. Simulat. 14 (6), 398–408.

Warm, J., Parasuraman, R., 1987. Vigilance: basic and applied research [Special issue]. Hum. Factors 29 (1), 623–624.

Warm, J.S., Parasuraman, R., Matthews, G., 2008. Vigilance requires hard mental work and is stressful. Hum. Factors 50 (3), 433–441. https://doi.org/10.1518/001872008X312152.

Willis, S.L., Tennstedt, S.L., Marsiske, M., Ball, K., Elias, J., Koepke, K.M., Morris, J.N., Rebok, G.W., Unverzagt, F.W., Stoddard, A.M., Wright, E., ACTIVE Study Group, for the, 2006. Long-term effects of cognitive training on everyday functional outcomes in older adults. JAMA 296 (23), 2805. https://doi.org/10.1001/jama.296.23.2805.