Search for a distressed swimmer in a dynamic, real-world environment

Victoria Laxton, Duncan Guest, Christina J. Howard &
David Crundall

Department of Psychology
Nottingham Trent University

Address for correspondence:

Victoria Laxton
Department of Psychology
Nottingham Trent University
50 Shakespeare Street
Nottingham NG1 4FQ

Email: Victoria.laxton@ntu.ac.uk

Alternative contact:

Prof. David Crundall
Department of Psychology
Nottingham Trent University
50 Shakespeare Street
Nottingham NG1 4FQ

Email: david.crundall@ntu.ac.uk

Search for a distressed swimmer in a dynamic, real-world environment

Laxton, V., Guest, D., Howard, C. J., & Crundall, D.

ABSTRACT

Visual search is increasingly being explored in dynamic, real-world environments. This includes swimming pools, where lifeguards have shown superior drowning detection in simulated environments. Here we explored if lifeguard superiority is observed in real-life scenes of a busy swimming pool. Experiment 1 required participants to identify real-life distressed swimmers in clips of busy pool activity via a touchscreen interface. Experiment 2 sought to replicate the first study, with the inclusion of eye-movement measures. Experiment 3 varied the methodology, using an occlusion method where clips were frozen and blurred shortly after target onset. The results demonstrated an experience effect, with lifeguards detecting distressed swimmers more often and faster than non-lifeguards. No clear differences were found in the eye-movements between groups; thus, we cannot conclude that the lifeguards' faster responses are due to better scanning strategies. The different methodological approaches revealed the occlusion method to have the larger effect size, supporting the growing evidence that occlusion may be a better test for dynamic target detection than traditional response-time tests. This research demonstrates that the clear lifeguard experience effect generalises to real-life pool environments with a large number of swimmers and real incidents. It could be used to inform lifeguard training tools and assessments.

Public Statement

This study used real video clips of swimmers in difficulty to explore lifeguard visual search using traditional reaction-time studies and a novel occlusion method. The results suggest that lifeguards have better detection of distressed swimmers compared to non-lifeguards, however this was not reflected in *overall* eye-movements measures. This research also shows support for an occlusion method for distinguishing between different groups visual processing in dynamic scenes for the detection of domain specific targets.

Search for a distressed swimmer in a dynamic, real-world environment

INTRODUCTION

Lifeguards are a vital component of swimming safety, both in pool and coastal settings. One of the key aspects of the role is to scan the water to look for swimmers showing signs of distress. Despite the importance of visual search in this job, little explicit training is provided in how to scan the scene. Similarly, there are only a limited number of studies that have investigated the role of visual search skill in the lifeguarding task. In order to provide better training for lifeguards, more research into the necessary visual search skills is required.

Although there is limited research considering lifeguards' visual search skills and performance, comparisons can be made to other areas in applied psychology research, such as airport security screening, driving or CCTV monitoring (Biggs & Mitroff, 2014, Crundall, 2016, Howard et al, 2013). In these contexts, visual search often involves the detection of a particular object (a target), amongst an array of other objects (distractors; Eckstein, 2011). Typically, participants with more experience or expertise within the context, tend to display superior visual search skills over more inexperienced participants. Lifeguard surveillance has many parallels with these contexts (Lanagan-Leitzel, Skow & Moore, 2015). For instance, in many real-world searches the target item is unspecified in appearance and location (Hout & Goldinger, 2015; Schmidt & Zelinsky, 2009). Drowning incidents are varied in terms of when and where they occur, how long they last, and their behavioural characteristics, which can range from stillness and submergence to thrashing and bobbing. Drowning incidents also tend to be very infrequent, which has been linked to "catastrophically low detection rates" in other contexts such as luggage screening, and has been termed the ultra-rare-item effect (Mitroff & Biggs, 2014, p284). Sustained attention in visual search tasks for rare targets can also lead a deterioration in vigilance. Vigilance decrements are often considered to be due

to a depletion of resources following a period of sustained attention (Oken, Salinsky & Elsas, 2006). In 'sit-and-stare' tasks where sustained attention is needed over a long period of time, such as lifeguarding or piloting an aeroplane, search performance usually decreases over time due to deterioration in vigilance.

When searching for unknown targets, general target templates can be used to apply knowledge of the common features a target may have (e.g. a swimmer thrashing at the surface), even though the specific details, such as the colour or location of the target, remain unknown (Schmidt & Zelinsky, 2009). In drowning, certain characteristics are displayed by swimmers in critical danger that may help guide search. Conscious (active) drowners will often display the instinctive drowning response (Pia, 1974) which may include thrashing the arms, submergence and re-emergence, and a vertical position in the water. Unconscious (passive) drowners, however, will often lie face down and motionless in the water, possibly due to head trauma or a medical incident (Doyle & Webber, 2016).

Contextual knowledge of a scene may also be used to guide search to likely target locations (e.g. pedestrians are found on pavements, birds are found in trees; Torralba, Oliva, Castelhano & Henderson, 2006). It is possible that lifeguards have built up a sufficient contextual knowledge to guide their surveillance of the pool or sea to areas of interest. For instance, knowledge of a local riptide may bias lifeguards' attention to that area in a beach scenario (Page, Bates, Long, Dawes & Tipton, 2011). However, there is a possibility that using prior knowledge of the scene (*scene priors*) could also negatively impact the search by guiding attention away from other valid areas of the scene.

While lifeguard surveillance shares similarities with visual search in other domains, there are limitations when making comparisons. Real-world visual searches that are reported in the

literature often focus on static images, which either have a target present throughout the search or no target object at all (Alexander & Zelinsky, 2012; Hess, Wismer, Bohil & Neider, 2015; Peelen & Kastner, 2014). In these traditional search tasks, participants are required to indicate if the target is present or absent, for example, detecting a dangerous item in an airport security scan (Biggs & Mitroff, 2014). It is unclear, however, how findings from static searches such as this, will apply to more dynamic scenes. In an abstract search task, Jardine and Moore (2015) found previously-efficient search performance for static arrays declined to little more than chance in dynamic displays. This was believed to be due to searchers constant need to update representations of scene information during dynamic tasks.

One further complication with realistic dynamic search, is that often a previous distractor must change state to be classed as a target (i.e. a swimmer may appear to be safe in the water initially, but then develop difficulties and turn into a drowning target). Given the evidence regarding memory for visual search (Peterson et al., 2001; cf. Horowitz & Wolfe, 1998), this makes the surveillance task even more challenging.

In laboratory-based studies the effects of dynamic stimuli that change in terms of movement, colour or orientation have been explored. Kunar and Watson (2011) investigated the impact of moving letters in a visual search array and found poor performance for moving stimuli in larger set sizes, compared to static searches. Target detection did improve in smaller set sizes for the moving letters however, becoming similar to the detection rates of static images. Target templates were also found to influence search outcomes. When the target template was non-specific (one of the 5 vowels in the alphabet), error rates were seen to be high; however, these were reduced when the template was more specific (e.g. if the target was the letter 'A').

In a more recent study, Muhl-Richardson, Godwin & Garner (2018) used changing colours of distractors to trigger a target onset. They suggest that a target in dynamic scenes is detected following onset often via direct fixation, however detection could also be primed by prediction from context (in this case, because the colours changed according to a pattern). In lifeguarding, context may also be used to prime detection of imminent drowning events, though this is likely to involve a more complex interplay of factors such as the relationship between characteristics of the swimmer (e.g. inexperienced) and the location (e.g. the inexperienced swimmer is floating into the deep end of the pool). This could also relate to research exploring contextual cueing, where faster target detection is reported when the context (configuration) of the display is the same and targets appear in consistent locations (Brockmore, Castelhano & Henderson, 2006; Brockmore & Henderson, 2006; Chen, 2000; Chun & Jiang, 1998).

Domain experience has been shown to improve detection in applied dynamic visual searches. In a study of CCTV operators, Stainer, Scott-Brown and Tatler (2013) found that experienced operators tended to use contextual knowledge to guide the search. For example, during night shifts experienced operators spent more time monitoring areas around night clubs, however in day shifts these environmental settings were rarely checked. Furthermore, Howard et al. (2013) found that trained CCTV operators appear to be able to identify suspicious events earlier than non-experts, with such operators making more consistent eye movements and judgements of suspiciousness than untrained observers. Differences in eye-movements between domain experts and novices have also been noted. For example, Savelsbergh, Williams, Kamp, and Ward (2002) found that experts used more efficient search strategies compared to novices in a study requiring participants to attend to

soccer goalkeepers. This included spending more time fixating target relevant areas, such as the kicking leg or the ball regions, particularly in the approach to foot-to-ball contact. Similarly, Mann, Williams, Ward, and Janelle (2007) note in a review of the sport literature that there are systematic difference*s* in visual search between expert and non-expert sportspeople. Experts were generally reported to make fewer fixations of longer durations, suggesting that experts are generally able to extract more task-relevant information in each fixation compared to novices.

Assessing lifeguards' visual search skills

Of the few studies that have actually investigated lifeguard visual search, the results support the effect of domain experience in target detection when lifeguards' performance is compared to that of control participants (Laxton & Crundall, 2018; Laxton, Crundall, Guest and Howard, submitted; Page et al., 2011). However, in the Laxton et al. (submitted) study it was found that there were no differences in eye-movements between experienced lifeguards and non-lifeguards (e.g. time to first fixate the target, processing time on target, etc.), even though the lifeguards still showed task performance superiority in detecting drowning targets. The authors acknowledged however that non-significant differences in the different eye-movement measures may have added up to potentially reflect a significant behavioural effect in response times.

A number of these studies used highly controlled stimuli, with simulated drownings (Laxton & Crundall, 2018; Page et al., 2011). In the case of Laxton and Crundall (2018), volunteers from a lifesaving club were recruited to swim in regimented fashion, occasionally staging prescribed drowning events, while Page et al. (2011) used low-fidelity, animated representations for swimmers in an artificial beach scene.

It is not clear however whether lifeguard experience in these highly contrived tasks translates to the detection of drownings in a real environment. There are a number of problems with such highly controlled stimuli that should be noted. First, the simulated drownings created by Laxton and Crundall (2018) were acted by lifeguards, based on what they expect to see when someone drowns (rather than what they might actually see). This potential bias may provide unconscious benefits for naïve lifeguards' detection of drowning targets in the subsequent assessment test. Second, there was a lack of variation in distractor and target behaviour, which may have led to drowning events being easier to detect. For example, the regimented swimming of the distractors might have increased the pop-out effect of drowning events. Such regimented swimming may also create a search environment that is less relevant to real world drowning incidents, with lap swimmers being less likely to get into trouble than children engaged in play activities. One final problem comes from the limited number of distractor swimmers in the Laxton and Crundall (2018) study (3, 6 or 9 swimmers). This may create a display that is very easy to parse for the lifeguards, and it is possible that with more complex displays, the benefit of experience may be less evident.

Problems associated with these highly-controlled stimuli could be overcome by using clips of naturalistic drownings. Such clips would ideally involve real drowning characteristics, which would overcome some of the problems associated with previous highly-controlled stimuli. Furthermore, naturalistic poolside footage would create a realistic setting in terms of the number and behaviours of distractor swimmers.

There are, however, several difficulties in obtaining naturalistic poolside and drowning footage. First, the infrequency of real drowning events does not make it feasible for the

footage to be recorded by the experimenter. Additionally, there are issues with obtaining permissions to film real people in the swimming pool and ethical issues around filming or using film of genuinely distressing incidents. To overcome these difficulties publicly-available video footage of real drowning incidents was sourced via the internet with permission from the original uploader. These videos are of individual incidents filmed from an American wave pool, over several summers, with lots of different target incidents. While these events have been filmed over different days and over a number of years, all clips are filmed from roughly the same camera location (with only minor variations), which provides some consistency over all clips. The main advantages of these real-event video clips is that they include high numbers of distractor swimmers (ranging from 23 to 89), who are engaging in naturalistic play-swimming behaviours which have greater potential overlap with the features associated with active drowning characteristics. We ensured that any drowning incident captured by the footage and shown to participants would not be distressing, by including only clips where the poolside lifeguards make a successful and timely rescue.

In the experiments presented here, accuracy and response times were measured while lifeguards and non-lifeguards searched for real incidents in the wave pool video clips. The aim of this research was to identify whether previous evidence for lifeguard superiority in artificial settings can be generalised to real scenarios. It was predicted that the lifeguards' experience and training would result in faster and more accurate responses to drowning-present trials compared to control participants. As the number of swimmers in the pool increases it is expected that response times and accuracy rates to detect the drowning victim will be degraded, though lifeguard experience should mitigate this effect.

EXPERIMENT 1

METHOD

Participants

Fifty participants were recruited to take part in the visual search experiment (mean age 24.6, 28 female). Twenty-five of these participants (mean age 23.0, 12 female) had completed compulsory qualifications in lifeguarding prior to testing and had a varying amount of experience in poolside lifeguard duties (4.49 years of lifeguarding experience on average). Lifeguards were recruited from local leisure centres in the Leicestershire and Nottinghamshire areas of the UK.

The remaining twenty-five participants (mean age 26.3, 16 females) had no lifeguarding experience. Non-lifeguard participants were an opportunistic sample from Nottingham Trent University, made up from a majority of postgraduate students and research assistants. Some participants were also recruited from the same leisure centre as the lifeguards (reception and gym staff).

Design

A 2 x 3 mixed design was employed, comparing experience (lifeguard vs. non-lifeguard) across set size (low vs. medium vs. high). There were 30 drowning-present trials that contained active (conscious) drowning targets. These trials were genuine incidents, caught on a pool-side camera, which required lifeguard intervention. Active drowning targets were classed as swimmers who were displaying distress behaviours or the instinctive drowning response (Doyle & Webber 2016; Pia, 1974). In addition to the 30 drowning-present trials, 15 non-drowning trials were also included. Of the 30 drowning present trials, ten trials

contained low numbers of swimmers (averaging 29.4, range 23-36), ten trials contained medium numbers of swimmers (averaging 46.8, range 39-52), and the remaining ten trials contained high numbers of swimmers (averaging 73.2, range 60-89). The 15 non-drowning trials were evenly split across set size, with 5 clips in each condition.

Accuracy and response times to detect the drowning target were recorded. Participants responded by making a touch-screen response on a laptop to indicate the location of a potential drowning incident. The target was not present at the start of the trial: while the particular swimmer is visible, they are yet to start drowning and therefore have not yet become the target. In the Laxton and Crundall (2018) study it was found that non-lifeguards were more likely to make a premature response to a trial. In that study only one response was allowed, and this led to a decrease in the amount of data for all non-lifeguards (potentially over-estimating the differences between the two groups). Allowing multiple responses to trials removes this systematic bias and provides a more conservative comparison across groups (Laxton et al., submitted).

In the current study, while multiple responses were allowed, a clip would terminate upon participants making a correct response. A feedback screen would then be shown, and the trial would then move onto the next clip. Correct responses were recorded if a response was made in the correct location on the screen and was made after drowning onset, or no response was made during a target-absent trial. Incorrect responses were recorded if no response was made in a drowning trial, a premature response was made that was not followed by a correct response, or a response following drowning onset was made in an incorrect location. It was not possible to respond too late to the drowning, as the clip ended following the drowning event (immediately prior to visible intervention from the poolside

lifeguard). In addition to response times, the screen coordinates of responses were recorded. Drowning onset of each clip was determined from the first signs of visible distress.

The experiment was created to run as a single, continuous, randomised block with feedback screens after each clip. A spatially-responsive window (invisible to participants) was placed around the drowning target, which covered an area measuring 250 x 140 pixels, in the horizontal and vertical axes respectively. This spatial window around the target accounted for .8% of the total screen area. This window was only active after the onset of the drowning and remained centred on the target. If the target moved on the screen, the spatial response window moved accordingly so that accurate locations of participants' responses were recorded. A touch response inside an active spatial window was considered a correct response to the incident. Before the presentation of each trial, a fixation cross was presented on the centre of the screen for 500ms.

Stimuli and Apparatus

Initial video footage, captured by a static poolside camera at an American wave pool, was accessed from YouTube with the uploader's permission to use for experimental stimuli[1]. Wave pool lifeguard rescue videos 1-42 were used in the experiment. The camera is stationed at the left-hand side of the pool at the deep end. The footage shows either a long shot of the pool, looking towards the shallow end or a zoomed in shot of just the deep end (see Figure 1). Big inflatable rubber rings can be seen in the pool as well as the swimmers.

Footage is completely naturalistic, with swimmers (mostly children) engaging in fun swim behaviour (e.g. chatting in a group with friends, riding on inflatable rings, swimming and playing). The drowning incidents are real swimmers in distress; however all video clips have

---

[1] Footage can be found at https://www.youtube.com/channel/UCnERyC7dwJwTvEyzYz6uxHw .

a real lifeguard performing a rescue in a timely manner (within the taught 10:20 second standard; 10 seconds to detect an incident and 20 seconds to make a response) and none of the rescued swimmers suffered any long term injury or distress from the incident. All distress incidents are either swimmers displaying the *instinctive drowning response* or weak swimmers showing obvious signs of distress and loss of floatation (Pia, 1974; Doyle & Webber, 2016). The drowning incidents were cut immediately prior to the local lifeguard entering the water to initiate a rescue (i.e. clips were cut before any evidence of an incident was provided by the actions of the lifeguard).

Forty-five clips were selected from the footage, evenly distributed across the varying set size levels. Fifteen clips contained no drowning incidents, with 5 in each set size condition. The clips varied in length, ranging between 9-35 seconds. Drownings occurred quasi-randomly within the trial, happening at some point after the first 5 seconds. The drowning incidents lasted between 2-19 seconds with clips ending immediately following the drowning. On average, a drowning incident in the low set size lasted 6.95 seconds, with 5.58 seconds and 6.11 seconds for the medium and high set sizes. A one-way ANOVA was used to explore the potential differences between the drowning durations over the 3 set sizes, but none were found ($F_{(2,29)} = .22$, $p = .8$).

In addition to the different number of swimmers in each clip, there were also differences with the location of the drowning target. For example, in some clips the drowning target was nearer to the foreground than targets on other trials and therefore made up more of the response window than targets further away from the camera. When 'near' and 'far' targets were compared directly, this factor produced the expected main effect with participants responding faster and more accurately to closer targets. We anticipated this

effect and therefore ensured that there were an equal number of 'near' and 'far' targets in each set size.

Trials were played without an audio track to avoid the participants hearing early responses from the real pool lifeguard raising the alarm to the drowning situation. The trials were run on a Lenova Yoga touch screen laptop, with a screen resolution of 2880 x 1620, running Psychopy. The trials were run in a randomised block, with a feedback screen after each trial. Participants could make localised responses on the touch screen of the laptop to indicate where a drowning incident occurred. Spatial response windows (invisible to participants) were centred on the drowning target and recorded correct localised responses.

Figure 1.   Four screen shots taken from the video stimuli with faces of swimmers blocked to protect their identity. Blocked effects were not included in the actual study.

Procedure

In order to recruit lifeguards, we arranged testing at local pools in two counties within the UK. The test was conducted in convenient locations within the pools, such as in a canteen area or in the poolside viewing area. Non-lifeguard participants were tested in similar conditions, using a common area of the university to ensure similar levels of potential distraction.

Participants were first asked to fill in a consent form and were then given instructions for the task. Participants were told the nature of the experiment before starting, including that they may see some distressed swimmers and that video clips are of real pool footage. The participants were also made aware that they may withdraw at any point during the experiment if they did not wish to continue. Before the main experiment began, an on-screen demographic questionnaire, and a touch-screen practice test, was presented. For the touch-screen practice test, participants were asked to touch all green circles that appeared on the screen and ignore any red circles. Following this, a practice drowning trial was presented. Participants were informed that they could make multiple responses to each clip, though they were discouraged from making excessive or random responses. The practice trial did not contain a drowning, therefore did not require the participants to respond. Participants were given correct or incorrect feedback for the practice trial and told there was no drowning. They then started the main experiment. All 45 trials were presented in a single, randomised block, with each clip preceded by a 500 ms fixation cross and followed by a feedback screen. After completion of the main block, participants were thanked for their time and fully debriefed. This experiment was conducted with approval from the University's ethical board and run in accordance with the British Psychological Society Guidelines.

RESULTS

Analysis for all dependent variables (accuracy, RT) was completed using a 3 x 2 mixed ANOVA with set size (low, medium and high) and group (lifeguards and non-lifeguards) as independent variables. If set size produced a significant main effect or was involved in a significant interaction subsequent planned comparisons were employed, comparing the low and medium set sizes, and comparing the medium and high set sizes. Where there was an

interaction the planned comparison between these set sizes included the experience factor in order to identify the locus of the interaction. Where significant interactions required further exploration, t-tests were used, with adjustments for multiple comparisons using the Bonferroni correction.

Catch trial responses

The response rates to non-drowning trials were analysed first. Although non-lifeguards successfully avoided responding to 78.9% of catch trials on average and lifeguards successfully avoided 71.2% of catch trials on average, this difference was not found to be significant (*t*(48) = 1.39, p = .17).

Behavioural responses

The trials that received correct responses were converted into percentages and subjected to a group x set size (2 x 3) mixed ANOVA. Trials with a drowning target were considered incorrectly responded to if no response was made following the onset of drowning activity or a response was made to an incorrect location.

The main effect of experience (*F*(1,48) = 12.2, *MSe* = 157.3, *p* < .05, η_p^2 = .20) demonstrated that lifeguard participants were more accurate at detecting the drowning swimmer than non-lifeguards (77.2% vs. 64.8%, respectively). The main effect of set size (*F*(2,96) = 50.0, *MSe* = 166.1, *p* < .001, η_p^2 = .51) was subjected to planned repeated contrasts which showed that the low set size was responded to less than the medium set size (F(1,48) = 4.2, MSe = 295.4, p < .05, η_p^2 = .08), and the medium set size was correctly responded to more often than the high set size (*F*(1,48) = 83.0, *MSe* = 358.4, *p* < .001, η_p^2 = .63) (75.8% vs. 80.8% vs. 56.4% for the low, medium and high set sizes respectively; see Figure 2). While the

interaction between experience group and set size looked promising, it was not significant ($F(2,96)$ = 2.3, *MSe* = 166.1, *p* = .103, η_p^2 = 005).
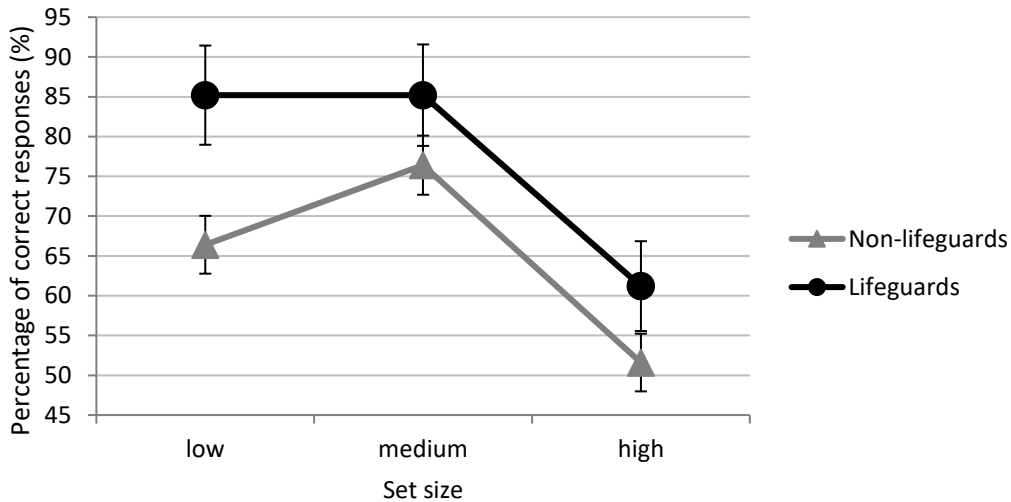
Figure 2. Mean percentage of correctly identified targets (with standard error bars)

A similar group x set size (2 x 3) mixed ANOVA was conducted for the response times of correctly identified targets only. Missing data for one participant was noted and this participant was removed from the following analysis. The main effect of experience ($F(1,47)$ = 8.6, *MSe* = 449285, *p* < .05, η_p^2 = .15) revealed that lifeguards were faster to respond to correctly identified drownings than non-lifeguards (3551 ms vs. 4113 ms, respectively).

The main effect of set size ($F(2,94)$ = 22.3, *MSe* = 737263, *p* < .001, η_p^2 = .32) when subjected to planned repeated contrasts noted that RTs in the medium set-size were shorter than the high set size ($F(1,47)$ = 57.7, *MSe* = 1006757, *p* < .001, η_p^2 = .55), though they did not differ to those produced by the low set size ($F(1, 47)$ = 1.13, *MSe* = 17464567, *p* = .29, η_p^2 = .02) (low: 3603 ms, medium: 3402 ms, high: 4490 ms; see Figure 3). The interaction between set size and group was not significant.
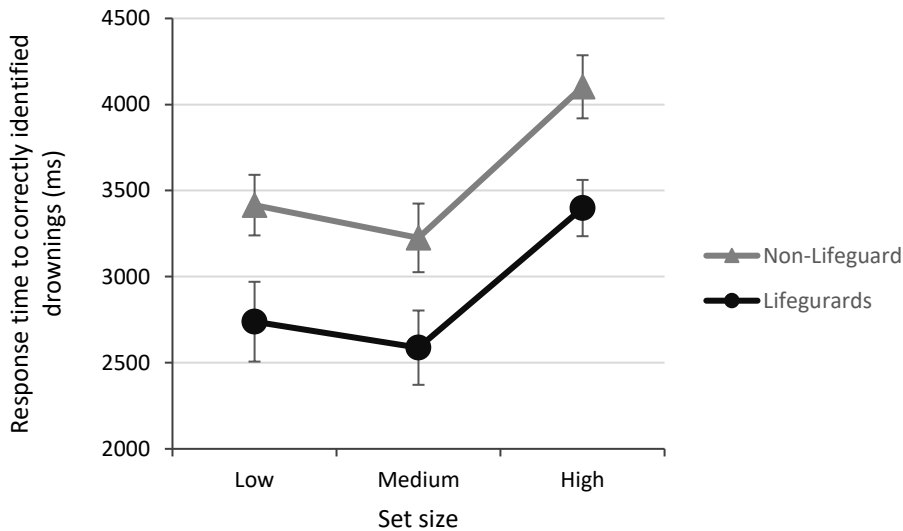
Figure 3. Mean response time to correctly identified drowning targets in ms (with standard error bars)

DISCUSSION

The results of this first experiment confirmed the superiority of lifeguard responses to real

drowning and distress. Overall, lifeguards were able to detect more drowning swimmers

than the non-lifeguard participants. A similar pattern was observed by Laxton and Crundall

(2018) with lifeguards in their experiment correctly identifying more of the simulated

drowning swimmers than non-lifeguards, supporting the idea that lifeguard experience

influences search in more complex real-world environments. This experience effect is also in

line with other types of surveillance-based visual search tasks in real world settings, where

individuals with more domain experience are found to have better search outcomes than

novices in both static and dynamic settings (Biggs & Mitoff, 2014; Curran et al., 2009).

It is interesting to note that the set size manipulation appeared to have little impact on

detection accuracy and RT when going from the low to medium set sizes, but led to a

significant deterioration in detection accuracy (of around 20%) for the high set-size stimuli

relative to the intermediate set-size stimuli. Although detection accuracy was reduced by

around 20%, the lifeguards still showed superior performance. It is probable that in the highest set size the scene was too cluttered to be able to search efficiently and this led to a different form of search strategy, or at least a less effective search strategy, resulting in a step-change in performance. Although the set sizes used in Laxton and Crundall (2018) are not directly comparable to this study, they also suggested a search strategy change at higher set sizes. They suggested that searchers potentially used a holistic strategy focussing on spaces between swimmers when only three or six people were in the pool, but when that increased to nine swimmers the strategy may have switched, encouraging fixations on individual swimmers. Despite step-change in performance in the high set-size condition, lifeguards still out-performed non-lifeguards. This indicates that, if search strategies did change from the medium to high set size, relative task accuracy across the groups did not depend on the search strategy. Overall, lifeguards were found to respond over half a second faster when correctly identifying drowning targets. These faster responses may potentially reflect lifeguards' faster recognition of drowning characteristics once the target is fixated.

One additional result to note is that the lifeguards were just as likely to make false alarm responses on catch trials as the non-lifeguard participants. Previous research has found that lifeguards are less likely to make a response in non-drowning trials (Laxton & Crundall, 2018); however, that results was found with relatively low set sizes and simulated drownings. It may be that these real drowning clips of highly cluttered swimming pools encourage a lower threshold for responding, resulting in more false positive responses among lifeguard participants. In real lifeguarding situations a lifeguard needs to make a quick decision to perform a rescue, assessing the situation to engage in an appropriate action or decide how best to proceed (White, 2017). To aid with this decision process

lifeguards are encouraged to use colloquial phrases such as 'when in doubt, check it out', or 'if you don't know, then go'. As a result, it may be that the realistic footage increases the number of false-alarm responses made by lifeguards, compared to the regimented lap-swimming of Laxton and Crundall (2018), effectively nullifying the effect found in the previous study.

In summary, experiment 1 verified that the lifeguard experience effect previously shown by Laxton and Crundall (2018) and Laxton et al. (submitted) using artificial stimuli is evident when using real drowning scenes. However, it provides little information on the processes underlying this effect. To examine this further Experiment 2 measured eye movements of lifeguards and non-lifeguards when watching these real drowning clips.

EXPERIMENT 2

Why do differences in visual search occur between lifeguards and non-lifeguards? There are both logical and theoretical reasons to believe that lifeguard superiority should be detected in visual search strategies, such as better scanning from exposure to pool environments or potential saliency of certain drowning types. Laxton et al. (submitted) explored this lifeguard superiority, but did not find any differences between lifeguard and non-lifeguard eye-movements following the onset of drowning events. However, they did note that small gains in time-to-first-fixate and target verification time added up to mirror the significant behavioural difference in response times between lifeguards and non-lifeguards.

It is possible that Laxton et al.'s null effects in regard to eye movements were in part due to the artificial nature of the stimuli. It is possible that more realistic stimuli (as used in the current study) might be better able to identify eye movement measures that underlie the performance benefit of lifeguards. This could be expected given that previous research into

real-world dynamic search tasks have found clear differences between experts and novices in eye-movements when carrying out surveillance tasks related to their domain expertise. For example, Bertram, Helle, Kaakinen and Svedström (2013) found that expert radiologists used saccades of shorter amplitude when detecting lymph nodes compared to a student control group. Furthermore, Konstantopoulos, Chapman and Crundall, (2010) found experienced drivers were quicker to fixate hazards and fixated safety-relevant stimuli for shorter amounts of time.

Experiment 2 therefore aimed to explore any differences in the eye-movements of lifeguard and non-lifeguards to the naturalistic clips used in Experiment 1. It was predicted that lifeguards would once again demonstrate a benefit in their behavioural responses to trials, and that this superiority would also be reflected in the eye-movement data, with the lifeguards fixating more of the drowning swimmers and fixating them earlier than non-lifeguards.

METHOD

Participants

Sixty-two participants (34 female) were recruited to take part (mean age of 21.7). Thirty-one of these participants (mean age 22.8, 7 females) had completed compulsory qualifications in lifeguarding prior to testing and had a varying amount of experience in poolside lifeguard duties (a mean of 2.5 years of lifeguarding experience). Two participants had completed compulsory lifeguarding qualifications but were noted to be working their first lifeguard shift on the day of testing. Lifeguards were recruited from advertisements on social media sites Linkedin, Twitter and Facebook, and were all from local pools in Nottinghamshire and Leicestershire.

The remaining thirty-one participants (mean age 20.4, 27 females) had no lifeguarding experience. Non-lifeguard participants were an opportunistic sample from Nottingham Trent University School of Social Sciences, made up from a majority of undergraduate students.

Design

The design was identical to that employed in Experiment 1, with the following exceptions. First, participants' eye movements were recorded while viewing the stimuli via a desktop eye tracker. Second, the mode of response changed, with participants no longer using a touch screen to identify the drowning swimmer. This change was necessary as the eye-tracker does not natively interface with a touch screen. Instead, participants were required to make a push button response if they identified a drowning swimmer (which is likely to increase Type 1 error rate compared to Experiment 1).

Fixations on targets were calculated by defining an area-of-interest (AOI) around the drowning target (no AOIs were used in catch trials). AOIs were only active following drowning onset and were invisible to participants. AOIs were identical to the spatial response windows in Experiment 1, averaging 2.5cm x 1.8cm in size (0.8% of the screen), and moving with the target where necessary

Apparatus and Stimuli

The stimuli were the same as those used in experiment 1. The experiment was presented on a Dell computer screen connected to an SMI RED500 eye tracker sampling at 500Hz. A saccade was defined as a change in gaze between two points with a minimum peak velocity threshold of 40°/s. A fixation was defined as a gaze with velocity below 40°/s (not a saccade) with minimum duration of 80 ms to allow brief re-orientation fixations to be included. The

trials ran in Experiment Centre as a randomised block. Before each new clip a fixation cross was shown, this would start the next trial when a participant fixated upon it for 500ms. Fixation windows were designed to replicate the response windows used in Experiment 1.

Procedure

In order to recruit lifeguards, the experimenter arranged testing sessions at various pools and leisure centres around the U.K., with a quiet office or side-room acting as the laboratory. Non-lifeguard participants were tested under similar conditions, using a small room within the university. Participants were given written instructions and asked to fill in a consent form and demographic questionnaire. Prior to the experiment, participants were made aware that they would be searching for any potentially drowning victims from a lifeguard's perspective. Participants were made aware that each drowning trial only contained one drowning incident, though they could make multiple responses if they thought they had made a false-alarm response. They were however discouraged from making excessive or random responses. Unlike Experiment 1, participants did not touch the screen to register a response (using the eye tracker precluded this). Instead, participants were told to respond via the zero key on the number pad of a standard keyboard. Once all instructions had been given, participants were given the opportunity to complete a practice trial, which was followed by a final opportunity to ask any remaining questions before the trials began. Once this was complete, eye tracking calibration took place, which required them to follow a moving cursor with their eyes while sat at 60 cm distance from the screen. Once the participant had been successfully calibrated to the eye tracker the test began. Upon finishing the test, the participants were fully debriefed and thanked for their time and participation. This research was conducted with approval obtained from Nottingham Trent

University ethics committee and run in accordance with British Psychological Society guidelines.

RESULTS

Catch trial responses

The response rates to the non-drowning trials were assessed first. On average, non-lifeguard participants avoided responding on 86.9% of catch trials, while lifeguards were ostensibly less successful with 78.3%, however this difference was not significant (t(60) = -1.87, p < .067).

Behavioural responses to drowning present trials

The percentage of trials with a drowning target that received correct responses is shown in Figure 4. Responses were considered incorrect if no response was made following the onset of drowning activity. The trials that received correct responses were converted into percentages and subjected to a group x set size (2 x 3) mixed ANOVA. One outlier, who responded to 80% of catch trials, was identified in the lifeguard group. The analysis was run with and without this participant, though the pattern of results was noted to remain the same. The following analysis excludes this individual.

A main effect of experience was found (F(1,59) = 19.8, *MSe* = 239.0, p < .001, η_p^2 = .25), with lifeguards successfully identifying 75.9% compared to the non-lifeguards identifying 58.3% of drowning targets. When the main effect of set size (F(2,118) = 47.9, *MSe* = 152.5, p < .001, η_p^2 = .45) was subjected to planned repeated contrasts it was noted that the low set size did not differ from the medium set size (F(1,59) = .31, *MSe* = 192.1 p = .58, η_p^2 = .01). However, the medium set size did differ from the high set size (F(1,59) = 64.3, *MSe* = 357.5,

$p < .001$, $\eta_p^2 = .52$), with more drowning targets being identified in the medium set size compared to the high set size (low: 72.9%, medium: 73.8%, high: 54.6%).

An interaction was noted between experience and set size ($F(2,118) = 4.2$, $MSe = 152.5$, $p < .05$, $\eta_p^2 = .07$). The repeated contrasts indicated that the interaction was evident between the medium and high set sizes, ($F(1,59) = 5.0$, $MSe = 357.5$, $p < .05$, $\eta_p^2 = .08$), such that the deterioration of performance between the medium and high set sizes was greater for the non-lifeguards.
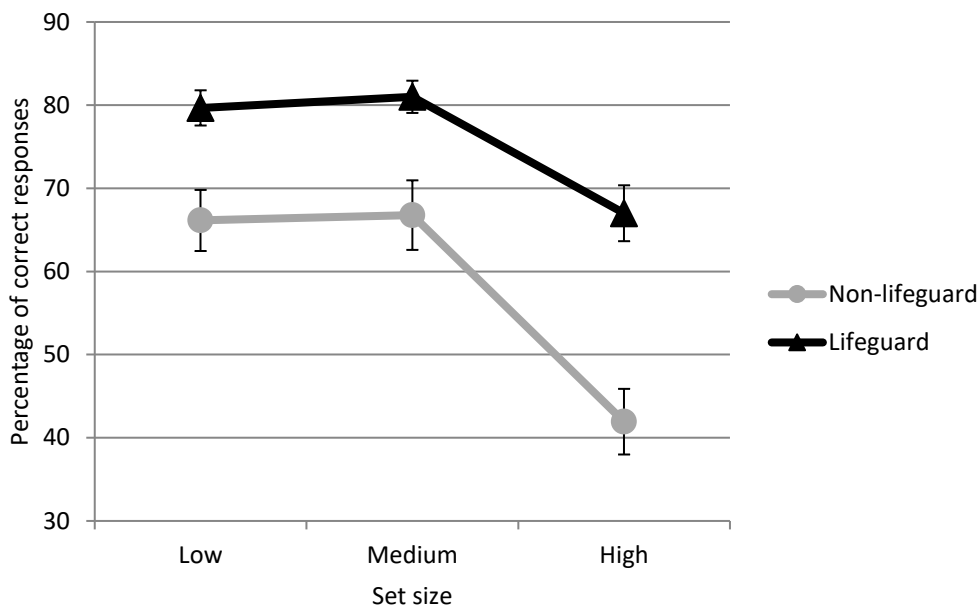

Figure 4. The mean percentages of trials containing a drowning target that were accurately responded to (with standard error bars)

The response times to correctly identify drowning targets (see Figure 5 were subjected to the same 2 x 3 ANOVA (experience x set size). Three empty cells for two non-lifeguards and one lifeguard were noted and these participants were removed from the following analysis. A main effect was found for experience ($F(1,57) = 5.9$, $MSe = 1387834$, $p < .05$, $\eta_p^2 = .09$),

which noted that lifeguards' responses during drowning event windows were faster than non-lifeguards by over 700 ms (3869 ms vs. 4615 ms, respectively).

The main effect of set size ($F(2,114)$ = 7.4, *MSe* = 1206982, $p < .05$, η_p^2 = .11) was subjected to planned repeated contrasts. This showed that the low set size differed from the medium set size ($F(1,57)$ = 12.1, *MSe* = 2309870, $p < .05$, η_p^2 = .18), with targets in the low set size being responded to faster than the medium set size. However, there was no difference between the medium and high set size ($F(1,57)$ = .02, *MSe* = 2225904, $p = .8$, η_p^2 = .00) (low: 3794 ms, medium: 4481 ms, high: 4450 ms). The interaction between experience and set size was not significant.
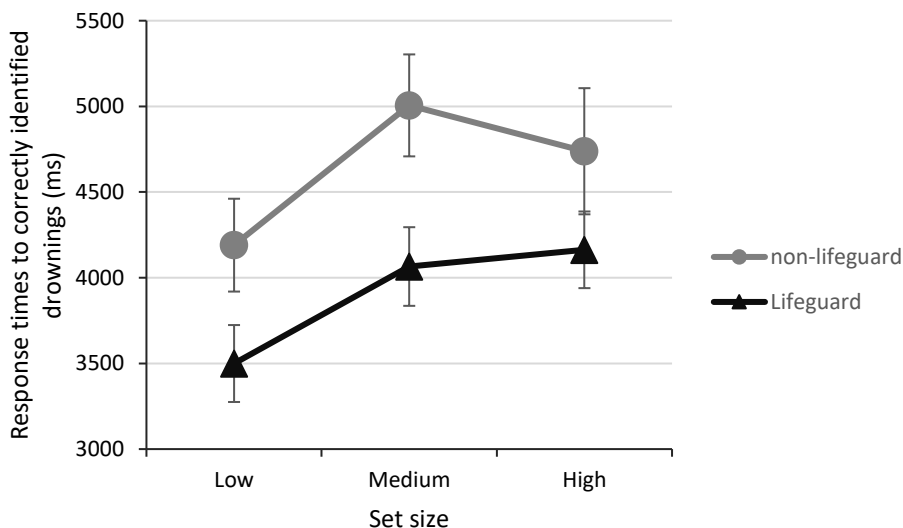

Figure 5. Mean response time to correctly identified drowning targets in ms (with standard error bars)

Eye movement data

Before analysing the eye tracking data, the tracking ratio for each participant was assessed, the tracking ratio was calculated by the eye-tracking software as the proportion of time that the eye-tracker recorded point of gaze coordinates over the entire task. All participants had

a good tracking ratio average for all trials (average 89.75%); therefore, no further participants were removed from the following analyses.

How many targets were fixated?
The percentage of targets that were fixated were subjected to an experience x set size (2 x 3) mixed ANOVA. Lifeguards fixated 79.6% of targets at some point following target onset, while controls fixated 85.8%. Compared to accuracy rates of 75.9% and 58.3% for lifeguards and controls, this demonstrates that mere fixation of the targets was not sufficient to produce an accurate response, particularly in the novice group. The difference in the number of targets fixated across the two groups was not significant (F(1,59) = 3.3, *MSe* = 196.0, p = .07, η_p^2 = .05). The main effect of set size was significant however (F(2,118) = 13.3, *MSe* = 96.0, p < .001, η_p^2 = .18). Planned repeated contrasts revealed that this effect was driven by the difference between the medium and high set sizes (F(1,59) = 11.8, *MSe*= 321.8, p < .05, η_p^2 = .17), with targets in the medium set size receiving more fixations than in the higher set size (84.5% vs. 76.7% respectively). The low set size was not significantly different from the medium set size. The interaction between experience and set size did not reach significance.

A check was also made to identify how many targets were actively fixated at the point in time when a successful response was made. The percentage of targets that were fixated at the time of a correct response was 68.2% and 50.4% for lifeguards and non-lifeguards, respectively (t(60) = -3.94, p < .001). Compared to accuracy rates of 75.9% and 58.3%, this suggests that approximately 8% of correct responses were made while the participant was not looking at the target. In these instances of non-fixation, participants had at least fixated the target once before moving their eyes elsewhere while making the response.

Time taken to first fixate the targets

The mean time (ms) to first fixate the target from the onset of drowning was subjected to a similar 2 x 3 ANOVA (experience x set size). The main effect of experience was not significant, with lifeguards taking an average of 2135 ms to fixate targets following onset compared to 2171 ms for non-lifeguards ($F(1,59)$ = .04, *MSe* = 471062, p = .8, η_p^2 = .001). The main effect of set size did reach significance ($F(2,118)$ = 6.4, MSe = 846830, p < .05, η_p^2 = .10). Planned repeated contrasts revealed there to be a difference between the low and medium set sizes ($F(1,59)$ = 6.5, *MSe* = 1588763, p < .05, η_p^2 = .10), and a difference between the medium and high set size ($F(1,59)$ = 13.6, *MSe* = 1432554, p < .001, η_p^2 = .19) (low: 2239 ms, medium: 1828 ms, high: 2394 ms. The interaction between experience and set size was not significant.

Dwell times on targets

Dwell times upon the target, as a percentage of the total duration of the drowning incident, were also subjected to a 2 x 3 ANOVA (experience x set size). A main effect for set size was noted ($F(2,118)$ = 99.5, *MSe* = 10.7, p < .001, η_p^2 = .63), with repeated contrasts showing that there is a significant difference between the low and medium set size ($F(1,59)$ = 60.8, *MSe* = 19.2, p < .001, η_p^2 = .51), and a difference between the intermediate and the largest set sizes ($F(1,59)$ = 40.9, *MSe* = 23.6, p < .001, η_p^2 = .41). Increased set size understandably reduced the amount of dwell on the target (with means of 20.8% vs. 16.4% vs. 12.4%, for low, medium and high set sizes, respectively). The main effect of experience did not reach significance ($F(1,59)$ = .79, *MSe* = 27.9, p = .37, η_p^2 = .01). The interaction between experience and set size did not reach significance.

Number of fixations made to target

The number of fixations made to drowning swimmers following onset were also analysed using an experience group x set size (2 x 3) mixed ANOVA. The main effect of set size ($F(2,118)$ = 42.3, *MSe* = 2.2, *p* < 0.001 η_p^2 = 0.42) was subjected to planned repeated contrasts. This revealed that the low set size differed from the medium set size ($F(1,59)$ = 23.2, *MSe* = 4.4, *p* < .001, η_p^2 = 0.28), and the medium set size differed from the high set size ($F(1,59)$ = 18.8, *MSe* = 4.3, *p* < .001, partial eta = 0.24; low: 8.0, medium: 6.7, high: 5.6).

The main effect for experience group ($F(1,59)$ = .13, *MSe* = 8.13, *p* = .718, η_p^2 = .00) and the interaction between experience group and set size ($F(1,118)$ = .07, *MSe* = 2.15, *p* = .927, η_p^2 = 0.00) were not significant.

Saccades

Mean saccadic amplitudes were calculated for each participant and compared in a mixed 2x3 ANOVA across group and set size. There was no difference between the two groups ($F(1,61)$ = .37, *MSe* = 33369, *p* = .543, η_p^2 = .01). There was also no difference between the set sizes and no significant interaction.

Two further measures were also explored: mean saccadic duration and saccade count. There was no difference between experience groups for saccadic average duration ($F(1,61)$ = .94, MSe = 9.91, *p* = .336, η_p^2 = .02) or saccade count ($F(1,61)$ = .45, MSe = 143252, *p* = .506, η_p^2 = .01).

Verification time

Further analysis was conducted looking at the target processing time calculated as the difference between participants' first fixation on a target and the subsequent response time. The time between the first fixation to the target and a behavioural response was

calculated to assess processing time. Responses where a target was not fixated were not included in the analysis. This was then subjected to a group x set size (2 x 3) mixed ANOVA.

The main effect of experience was not significant ($F(1,59)$ = 2.97, *MSe* = 2279051, *p* < .09, η_p^2= .48), despite the lifeguards having a processing time of 1913 ms compared to 2580 ms for the non-lifeguards. The main effect of set size ($F(2,118)$ = 6.24, *MSe* = 5511904, *p* < .05, η_p^2= .10) was driven by lower verification times in the low set size compared to the medium set size ($F(1,59)$ = 16.14, *MSe* = 3660661, *p* < .001, η_p^2= .22). The medium set size did not differ from the high set size ($F(1,59)$ = .80, *MSe* = 6011949, *p* = .37, η_p^2= .01) (low: 1684, medium: 2668, high: 2387). The interaction between experience and set size was not significant.

DISCUSSION

The results of the second experiment have further confirmed the predicted superiority of lifeguard responses to real drowning and distress and are in line with those of Experiment 1. Lifeguards correctly identified more drowning and distressed swimmers than the non-lifeguards. This superiority was also reflected in the response times to drowning and distressed swimmers, with lifeguards making responses that were over 700ms faster than non-lifeguards. These metrics of lifeguard superiority were not however reflected in the patterns of significant results derived from the eye movement measures.

It should be noted that the increased accuracy responses in the medium set-size that has been noted in the previous experiment was not present in the behavioural responses of the current experiment. Although there was an interaction between set size and experience in the accuracy of drowning detection in Experiment 2, this was only apparent in the shift from

a medium to high set size, with non-lifeguards' performance more negatively affected by the additional swimmers. This result is more in accordance with the well-known set size effect found in traditional controlled laboratory style experiments, which show degradations in performance as set size increases (e.g. Snodgrass & Townsend, 1980; Wolfe, Alaoui-Soce & Schill, 2017; Wu & Wolfe, 2016). This result differs from Experiment 1, which showed some set size effects where the medium set size evoked the best performance.

Why did the eye movement data not reflect the performance differences between lifeguards and non-lifeguards? In other naturalistic visual search tasks it has been demonstrated that having experience in a certain domain can influence eye movements during search, with experts being faster to fixate targets, make shorter fixations, and make fewer re-visitations than novices or people with no experience at all (Borowsky & Oron-Gilad, 2013; Konstantopoulos, Chapman & Crundall, 2010). While it might seem surprising to not find evidence of eye movement differences in lifeguards, such lack of effects has been noted by others in this domain. For instance, we previously found no overall difference in eye-movements between experienced lifeguards and non-lifeguards, even though the lifeguards still showed task performance superiority in detecting drowning targets (Laxton et al., submitted). However, we suggested that two non-significant eye movement measures in that study could potentially add up to reflect the significant behavioural response. Laxton et al.s', (submitted) lifeguards were 356 ms faster than non-lifeguards at first fixating the target, and their processing time on the target following first fixation was 395 ms shorter. Though not significant on their own, together these measures may add up to provide a potential speed advantage of 751 ms, which was close to the 720 ms response time advantage demonstrated by the lifeguards in that study.

Might a similar effect have occurred in the current study? Our current data show a non-significant advantage in verification time on target of 667 ms in favour of lifeguards (*p* =.09), though the lifeguard advantage in the time to first fixation the target was neglible (36 ms). Nonetheless, together these two non-significant effects add up measures still approximate the 746 ms response time advantage that demonstrated by the lifeguards.

The overall lack of eye-movement differences between the lifeguards and the non-lifeguards suggests that lifeguards have superior performance outcomes that are not necessarily reflected in changes in specific eye movement metrics. Instead, the higher rates of detection in the lifeguards may be due to other underlying cognitive mechanisms, such as a better ability to track multiple objects (e.g. following swimmers around a pool). There is precedence for this: Faubert (2013) found that professional athletes performed better at tracking multiple objects with attention in 3-D space.

Alternatively, processing efficiencies may yet underlie the lifeguard advantage. While one could argue that faster processing should be reflected in shorter gaze durations on the target (which did not reach significance in experiment 2), a range of post-processing activities may lead to fixations tarrying on targets even after identification. Certainly, the length of gaze on the target came closest to providing a significant explanation for the response time advantage.

Despite the success of experiment 2 replicating the behavioural results of the first experiment, one problem with the design was the use of a temporal scoring window to identify correct responses: without a spatial component to the behavioural response we cannot be sure that they were responding to the correct target, even if they responded at the correct time. This could have inflated performance (though Experiment 2 accuracy rates

were actually lower than those recorded in Experiment 1), and introduced a range of post-perceptual decision-making biases (discussed more in the following section.) A third experiment using these real drowning clips was therefore undertaken to further explore the superiority effect of lifeguard participants. This experiment employed an occlusion technique in which the video is stopped and overlaid by a still frame which is blurred out to prevent further extraction of detail from the scene. This was done to test if information can be extracted from the scene within a couple of seconds following drowning onset and if drownings can still be accurately located.

EXPERIMENT 3

Experiments 1 and 2 have shown an effect of superiority in lifeguards' ability to detect a drowning target in a visual search task containing real drowning incidents. However, these two experiments focussed on detection of drowning events, requiring participants to make a response within a temporal scoring window. These scoring windows were dependent on the length of the incident and ranged from 2-19 seconds. In cases where the drowning window is quite long, it is likely that even the non-lifeguards will eventually spot the drowning target. This may reduce the sensitivity in measuring lifeguard superiority in regard to the number of drowning targets detected.

Response times may also suffer from post-perceptual decision-making biases that are linked to an expert's threshold for detecting a target. For instance, while one might expect expert drivers to respond faster than novices to a hazard in a video clip of driving, evidence suggests that highly experienced drivers may have a higher threshold for acknowledging something to be hazardous (Crundall, Chapman, Phelps & Underwood, 2003, Crundall &

Kroll, 2018). Thus, expert drivers (or lifeguards, in this case) may spot a hazard/drowning target sooner than a non-expert participant, but then delay responding until a higher threshold of evidence is met. If this occurs on a small percentage of trials this may underestimate lifeguard superiority as measured by response times. It is well established that differences in speed/accuracy trade-offs between experimental conditions can limit the interpretation of single point response times (e.g., Dosher, Han & Lu 2010, Guest & Lamberts, 2011; McElree & Carrasco, 1999; Wickelgren, 1977).

One alternative to relying on potentially-confounded response times is to limit the stimulus exposure duration. In the hazard awareness literature an occlusion method based on the Situation Awareness Global Assessment Technique is being increasingly used (SAGAT; Endsley, 2017). In this occlusion method, a video clip is stopped at the point where a target event has just started, and a question is asked about what is currently happening or about to happen. Occlusion tasks in visual search are believed to isolate the predictive element in domain-specialist search and mitigate the problem of criterion bias inherent in responses times (Pradhan & Crundall, 2017).

These prediction tasks have recently been explored in driving research and have been found to be robust tests for discriminating between novice and highly experienced drivers (Castro et al., 2014; Crundall, 2016; Lim, Sheppard & Crundall 2014; Ventsislavova et al., 2019). For example, Crundall and Kroll (2018) compared a response-time measure of hazard perception in fire-appliance drivers to their scores on an occlusion task. The latter test was found to better differentiate between high-risk and low-risk drivers even though the stimuli were identical (barring the occlusion). This was also recently applied to CCTV footage. Crundall and Eyre-Jackson (2015) reported that expert police officers were better than

control participants at identifying the type of crime that is about to be committed in CCTV footage that is occluded at the onset of the criminal activity.

In Experiment 3, the same stimuli were used as in Experiment 1-2, but an occlusion method was employed. This method was adapted to enable the overall gist of the scene to remain visible to participants after the occlusion event. To do this, instead of replacing the video with a blank screen, it was replaced with a blurred, still frame of the video. The blurred screen allowed participants to guide their touchscreen response to a swimmer whom they may have identified as the target prior to occlusion. Crucially, the blurred screen prevented the extraction of any post-occlusion evidence for target identification.

This task enabled us to explore the effect of experience whilst controlling for differences in response criterion between lifeguards and non-lifeguards. Occluding just after the onset of the drowning behaviour event also reduces the possibility of non-lifeguard participants 'stumbling' across the target during the otherwise relatively lengthy period that the target is available. By removing accidental hits from the performance of non-lifeguards, this may increase the differences noted between the groups.

Median response times from Experiment 1 were used to determine when our clips should be occluded in Experiment 3. It was anticipated that this methodology would elicit a greater difference between lifeguards and non-lifeguards (reflected in effect size), with the non-lifeguard participants detecting fewer drownings than lifeguards, and response accuracy decreasing with increasing set size. Furthermore, we predict lifeguard superiority should be more apparent at higher set sizes.

METHOD

Participants

Fifty participants were recruited to take part in a visual search experiment using real drowning incident videos (mean age 23.2, 29 female). Twenty-five of these participants (mean age 24.3, 9 female) had completed compulsory qualifications in lifeguarding prior to testing and had a varying amount of experience in poolside lifeguard duties (4.24 years of lifeguarding experience on average). Lifeguards were recruited from local leisure centres and recreational parks in the East Midlands. The remaining twenty-five participants (mean age 22.0, 20 females) had no lifeguarding experience. Non-lifeguard participants were an opportunistic sample, recruited mainly from a university population.

Design

The same design was used as that of Experiment 1, comparing two experience groups (lifeguard vs. non-lifeguard) across set size (low vs. medium vs. high), in a 2 x 3 mixed design. The median response times of participants' responses to each clip from Experiment 1 were used to create cut points for the occlusion for each individual clip. At the median response point a blurred freeze frame was presented, with participants required to either touch the location where they thought a distressed swimmer was, or touch a 'No drowning' response box located in the bottom right corner of the screen. Occlusion points for the drowning-absent clips were randomly selected between 9 and 18 seconds.

Accuracy of responses was recorded via the same spatial response windows used in Experiment 1 (though these windows were static in the current experiment and only applied to the blurred occlusion screen). Correct responses were noted if a drowning swimmer was correctly identified, or if the trial was correctly identified as a no-drowning trial. If a

response was given outside of the spatial response window, then an incorrect response was noted.

Apparatus and Stimuli

The video clips used in Experiment 3 were the same as those used in Experiment 1. The 'no-drowning' response box was placed in the bottom-right corner with its own spatial response window (as seen in Figure 6. No swimmers were occluded by the 'no drowning' response box, with the box covering either the end of the pool or a section of pool that had been roped off.

The decision to use a blurred image rather than a complete occlusion was based on the need for participants to retain a spatial framework for making their touch-screen response. If the screen completely occluded, it would be possible that inaccuracies in reaching to give a touch response might increase due to memory errors for location, or simply the lack of a stimulus to reach toward. The blurred screen provides visual anchors to guide reaching behaviours, while still removing high-frequency information that would allow drowning detection via post-occlusion identification.

As in Experiment 1, a Lenovo Yoga touch screen laptop was used, with a screen resolution of 2880 x 1620, running Psychopy. The trials were run in a randomised block, with a feedback screen after each trial. Participants were able to make localised responses on the touch screen of the laptop.
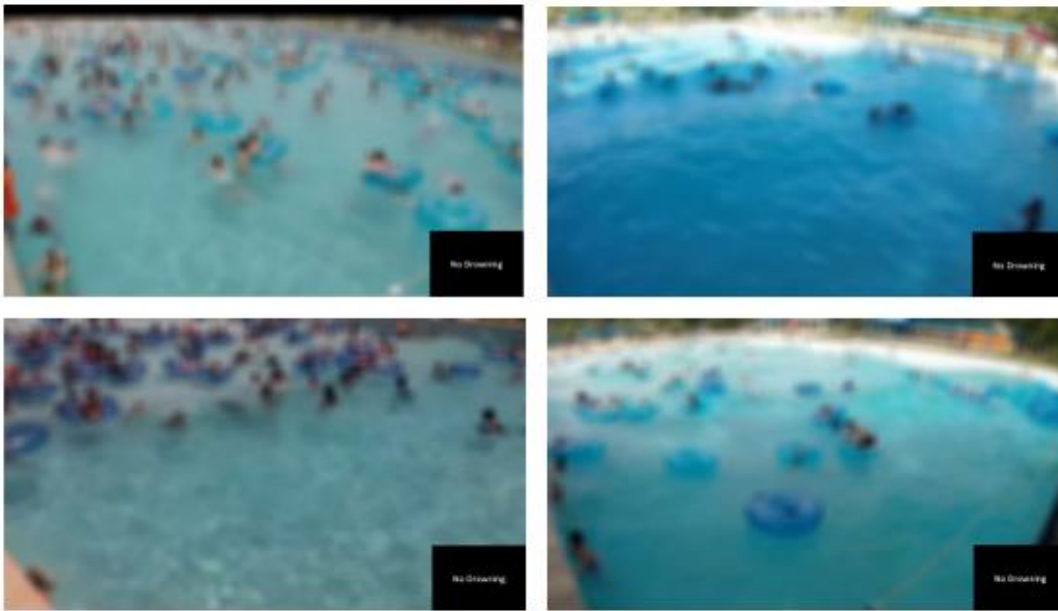
Figure 6. A selection of screenshots for various set sizes of the occlusion screens used in Experiment 3. The 'No Drowning' response box can be seen in the left-bottom corner of the images.

Procedure

The procedure was the same as that used in Experiment 1, with the exception that participants responded via touch screen once each clip had ended. Whereas Experiment 1 considered 'no-drowning' responses to be reflected in an absence of responses during the clip, Experiment 3 required participants to make a conscious 'no-drowning' response by selecting the box in the bottom-right corner.

RESULTS

Catch trial responses

The response rates to no-drowning trials were assessed first. On average, non-lifeguard participants successfully avoided making an incorrect response to 71.2% of trials, while the lifeguard participants successfully avoided making a response to 74.1% of trials. There was no difference in the number of trials successfully avoided between controls and lifeguards ($t(48) = .056$, $p = .09$).

Signal detection analysis

The design of this current experiment allows for the calculation of simple signal detection (STD) measures *d'* (sensitivity) and *c* (criterion), as the decision regarding the presence or absence of a drowning target occurred at the same point in the clip (post-occlusion). A hit was recorded when participants correctly located the drowning target. A false alarm was recorded when participants made a drowning-present response in drowning-absent trials. There was a significant difference in *d'* across the two groups with lifeguards showing greater sensitivity than non-lifeguards (1.09 vs. .47, *t*(48) = -2.67, *p* < .05). No difference was noted between the criterion scores of lifeguards and non-lifeguards (-.52 vs. -.6, *t*(48) = -0.56, p = .581).

Behavioural responses

Correct responses to drowning present trials were then assessed. Trials with a drowning target were considered incorrectly responded to if a response was made to an incorrect location, or a 'no drowning' response was made. The trials that received a correct response were then converted into percentages and subjected to a group x set size (2 x 3) mixed ANOVA. There were no outliers in the data and no participants were removed.

A main effect of group (*F*(1,48) = 17.7, *MSe* = 256.6, *p* < .001, η_p^2 = .43) revealed that lifeguards were more successful in correctly identifying the drowning swimmer than non-lifeguards (63.5% vs. 44.4% respectively). The main effect of set size (*F*(2,96) = 33.4, *MSe* = 198.1, *p* < .001, η_p^2 = .41) was subjected to planned repeated contrasts which found no difference between the low and medium responses (*F*(1,48) = 1.4, *MSe* = 401.5, *p* = .2, η_p^2 = .03), though the medium set size differed from the high (*F*(1,46) = 31.8, *MSe* = 510.2, *p* < .001, ηp2 = .40; low: 61.4%, Medium: 58.0%, high: 40.1%).

Despite an apparent trend towards an interaction between experience and set size (see Figure 7a), the effect did not reach conventional levels of significance ($F(2,96) = 2.9$, *MSe* = 198.1, *p* = .058).
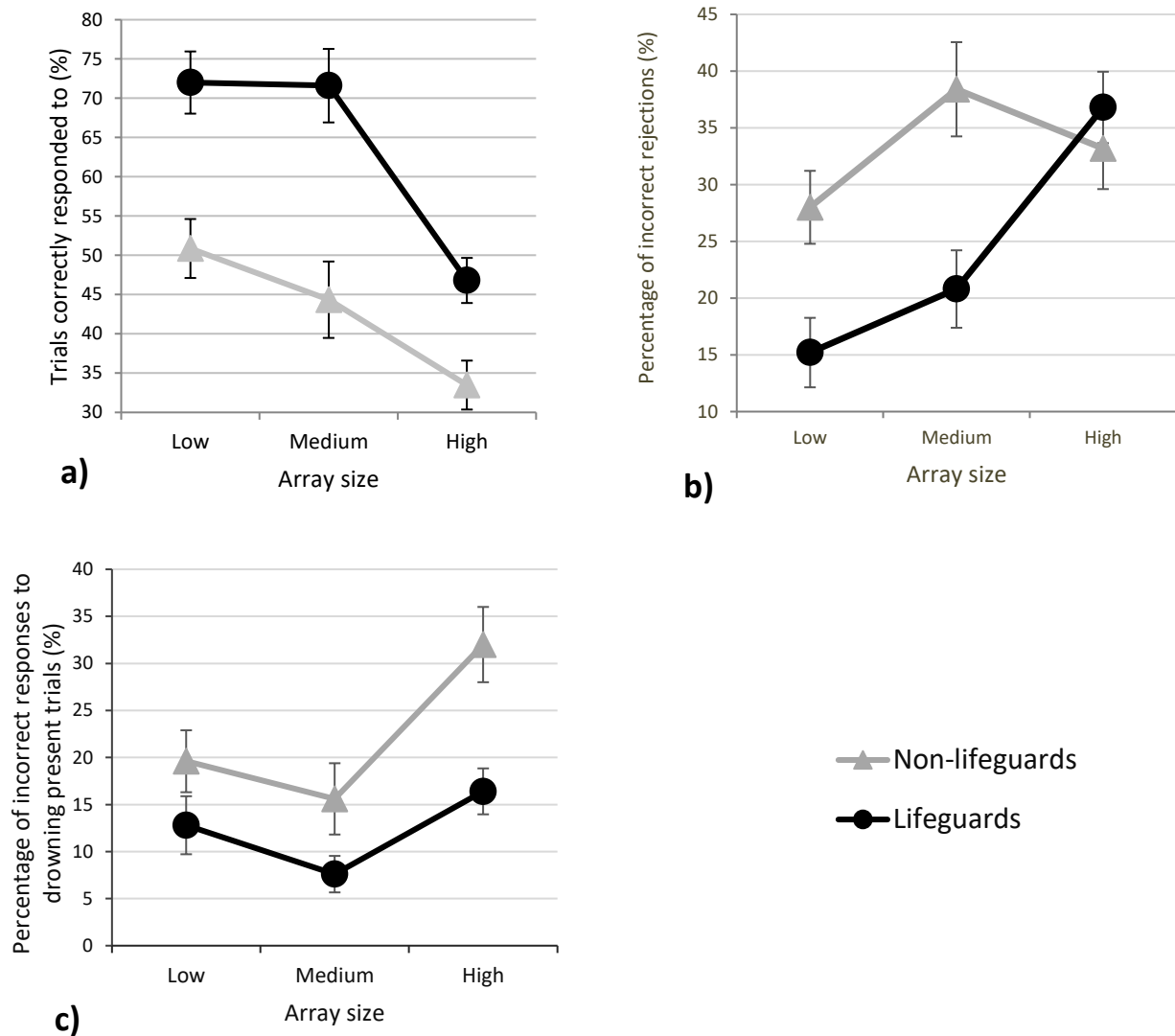

Figure 7. Mean percentages of a) correctly identified targets; b) no drowning responses to drowning present trials; c) incorrect location responses to drowning present trials (with standard error bars)

A similar group x set size (2 x 3) mixed ANOVA was conducted for the percentage of no-drowning responses recorded during drowning-present trials (i.e. incorrect rejections; Figure 7b). A main effect of group ($F(1,48) = 4.5$, *MSe* = 188.2, *p* < .05, η_p^2 = .09) revealed that lifeguard participants made fewer incorrect 'no-drowning' responses compared to non-

lifeguard participants (24.9% vs. 33.2% respectively). The main effect of set size (F(2,96) = 12.9, MSe = 171.0, p < .001, η_p^2 = .21) when subjected to planned repeated contrasts revealed the low set size differed from the medium set size (*F*(1,48) = 9.8, *MSe* = 327.3, *p* < .05, η_p^2 = .17), but the medium did not differ from the high set size (*F*(1,48) = 2.8, *MSe* = 480.0, *p* = .1, η_p^2 = .06) (low: 22.0%, medium: 30.0%, high: 35.2%).

An interaction between experience group and set size was noted (*F*(2,96) = 8.7, *MSe* = 170.9, *p* < .001, η_p^2 = .16). The repeated contrasts identified that experience interacted with set size only between the medium and high condition (*F*(1,48) = 11.3, *MSe*= 480.0, *p* < .05, η_p^2 = .19). Figure 7b appears to show that this was driven by an increase in the number of incorrect 'no-drowning' responses made by lifeguards in the high set size, where their responses become ostensibly indistinguishable from those of the non-lifeguards. Post hoc Bonferroni adjusted t-tests support this interpretation, with lifeguards making fewer no-drowning responses than non-lifeguards in the low and medium set sizes (low: *t*(48) = 2.6, *p* < .016; medium: *t*(48) = 3.1, *p* < .016). However there was no difference between the lifeguard and non-lifeguards responses in the high set size (*t*(48) = -.83, *p* = .41).

Incorrect location responses to drowning trials were also converted into a percentage and subjected to a group x set size (2 x 3) mixed ANOVA (see Figure 7c). The main effect of group (*F*(1,48) = 7.04, *MSe* = 173.3, *p* < .05, η_p^2 = .14) revealed that lifeguards made fewer incorrect responses than non-lifeguards (12.3% vs. 22.4% respectively). The main effect of set size (*F*(2,96) = 17.03, *MSe* = 117.7, *p* < .001, η_p^2 = .27) when subjected to planned repeated contrasts revealed that the low set size differed from the medium (*F*(1,48) = 5.0, *MSe* = 213, *p* < .05, η_p^2 = .09) and the medium set size differed from the high (*F*(1,48) = 28.4,

MSe = 280.0, *p* < .001, η_p^2 = .37; low: 16.2%, medium: 11.6%, high: 24.2%). The interaction between group and set size failed to reach significance.

DISCUSSION

The results of Experiment 3 were similar to Experiments 1 and 2. Lifeguard participants in Experiment 3 were found to make more correct responses to drowning-present trials compared to non-lifeguard participants. The superior performance of lifeguards is also in line with the findings of Laxton and Crundall (2018) and Laxton, et al. (submitted), where lifeguards detected more simulated drowning swimmers than non-lifeguards during searches of dynamic (but staged) pool scenes. In conjunction with the results from Experiments 1 and 2, this consistent experiential effect demonstrates that lifeguard drowning detection performance translates from the simulated and highly controlled task in Laxton and Crundall (2018) to the naturalistic stimuli used in the current studies.

One potentially important finding to note is the difference between the methodologies of Experiments 1 and 3. Experiment 3 produced a larger effect size for the difference between lifeguards and non-lifeguards than the response time study employed in Experiment 1 (η_p^2 = .41 vs .2). The greater effect size in Experiment 3 may be due to the mitigation of criterion bias by removing response time measures, which may have biased the responses in Experiment 1. Furthermore, some longer clips may elicit more responses than the shorter clips, as participants who may not initially see the drowning target in the early stages may happen across the target at a later point in the clip. During searches of longer clips, participants would have more time to look through the search display after drowning onset rather than relying on early cues, which could mask the benefits of experience in response

accuracy. The greater effect for the occlusion task is in line with other research, which has demonstrated that occlusion tasks are a more robust way of assessing expertise in these judgements as the ambiguous response time windows are removed (Pradhan & Crundall, 2017).

In the current experiment, lifeguard participants were found to make fewer incorrect 'no drowning' responses during drowning present trials (incorrect rejections) and fewer incorrect responses to drowning present trials, where a non-drowning swimmer was incorrectly identified as the drowning target. This was particularly apparent in the low and medium set size for 'no drowning' responses during drowning present trials. This demonstrates that the lifeguards are able to better recognise and respond to drowning signals compared to non-lifeguards, though extremely high set sizes may still pose a problem.

GENERAL DISCUSSION

Three experiments have consistently shown lifeguards to have superior detection of drowning swimmers in a naturalistic search task. The lifeguards detected more drowning swimmers across all experiments and had faster response times in Experiments 1 and 2. The superiority of lifeguard search found in all three experiments also fits with previous research that has explored lifeguard's drowning detection in simulated, naturalistic drowning scenes (Laxton and Crundall, 2018). However, no significant differences were found between eye-movement measures for lifeguards and non-lifeguards, suggesting that the greater performance of lifeguards may instead result from differences in underlying cognitive mechanisms, rather than a superior visual search strategy *per se*. As a caveat to this conclusion however, it remains possible that non-significant lifeguard advantages in the

time taken to first fixate a target, and the time between first fixation and subsequent response, add up to produce the significant response time superiority for this group.

One difference between this current set of experiments and previous similar work (Laxton & Crundall, 2018; Laxton et al., submitted) is the more complex, realistic stimuli used to explore lifeguard drowning detection. It could be argued that the pool footage used in Laxton and Crundall (2018), is less challenging than the realistic stimuli used in the current studies, and places fewer demands on visual search, leading to an overestimation of drowning-detection performance. This may have especially benefitted the lifeguards. The findings of the current experiments however suggest that lifeguards remain superior in detecting drowning events when tested with real-world drowning incidents in a much more complex visual environment. When designing a test of lifeguards' visual search performance, it is difficult to obtain video footage of real drowning incidents, however the consistency of the results across the current experiments (using real drowning incidents) and those involving simulated incidents (Laxton & Crundall, 2018) suggests that future tests could use simulated drownings with the expectation that evoked results will generalise to real-world incidents.

Despite this experiential effect being consistent across the behavioural data of all three experiments, no clear explanation for this effect was found in the eye-movement data in the measures analysed. The lack of group differences in the measures analysed in the current research does not suggest that lifeguards have a better visual search strategy. Therefore, we must consider that any performance advantages the lifeguard participants demonstrate might instead be due to how they process the visual information taken in during the scanning. Although there is the potential for non-significant gains in verification time and

time to first fixate the target to add up to the response-time difference, the primary driver behind this post-hoc rationalisation is the relatively large (though still non-significant) difference in the verification time between the groups. Faster processing of drowning and distress behaviours could be a result of exposure to such behaviours, both in training and in real-life incidents. While faster processing should be reflected in shorter gaze durations on targets (Gegenfurtner, Lehtinen & Säljö, 2011), it is possible that our lifeguards delayed a response in some trials while remaining fixated on the target. More research is needed to explore the exact nature of the underlying lifeguard advantage, and to assess the true extent that this might be reflected in measures of gaze durations on targets.

It is interesting that a clearer expertise effect was not found in the overall eye-movements when other types of applied dynamic visual search tasks have found clear differences between domain experts and novices. In tasks such as driving or sports, experts are found to make faster eye-movements to target relevant areas (Howard et al., 2010), have fewer fixations of longer duration (Savelsbergh et al., 2002) and make fewer re-visitations to targets (Konstantopoulos et al., 2010). Why might the task of drowning detection not follow a similar pattern to other dynamic tasks? The dynamic search task of the lifeguard differs slightly to other types of dynamic visual search in that there are many more swimmers in the visual field than there would be cars/hazards on a road or players in a sports game for example. This creates a highly cluttered environment and makes target detection difficult at times. For example, fixations may land on the target, but processing of the visual information may be constrained or incomplete due to the urge to maintain an active search. This could potentially relate to the static image search of airport security screening, where experts are required to detect dangerous items in often cluttered bags (McCarly, Kramer &

Wickens, 2004). Being able to see through the clutter and quickly process the items, or the swimmers' behaviours in the case of lifeguarding, becomes important for making such high-stakes decisions.

The clear and consistent behavioural superiority of lifeguards echoes and extends the findings in other domains, including search in both static and dynamic real-world environments (Biggs & Mitroff, 2014; Evans et al., 2011; Konstantopoulos, Chapman & Crundall, 2010; Reingold, Charness, Pomplun & Stampe, 2001). In addition to demonstrating a robust experiential effect in a novel domain, some current findings contrast with those from other domains. For example, why are fixations on hazards typically shorter for experienced drivers (Chapman and Underwood, 1998), though in the current data (and in Laxton et al., submitted) our experiential effect is always tantalisingly out of reach? Perhaps there are differences between these two domains which might explain why it is easier to find this effect in driving. For instance, a car driver is typically advancing through the visual scene. This provides added impetus to maintain an active search. Even when a primary hazard is detected, the safest drivers will be concerned with identifying potential secondary hazards. Furthermore they may need to continue to search the scene in order to ensure that their response to a hazard ahead will not negatively affect other road users (e.g. a decision to swerve around a hazard should be dependent on the presence of vehicles in the adjacent lane). In lifeguarding however, the lifeguard is stationary. There is no behavioural urgency to resume an active search once the target has been found, which may result in longer fixations than are necessary, and (despite mantras such as 'if in doubt, check it out') a tendency to delay responses until greater evidence is gathered. This could be similar to the expert search of airport baggage screening (Donnely et al., 2019). For airport baggage

screeners, suitcase objects will always be stationary and will always be present from the start of the search. Therefore, once identified these targets can be accepted and that bag can be taken for further investigations. However, for lifeguards, there is the potential for any swimmers to change into the 'target' and drowning incidents to develop over time, therefore longer fixations and verification times would be beneficial.

Limitations and future considerations

All three experiments employed video stimuli without an audio track to avoid the participants hearing early responses from real lifeguards stationed at the pool. This potentially results in the video clips lacking some external validity. It could be possible that lifeguards are primed to detect particular audio cues (e.g. shouts for help from nearby swimmers, changes in the pattern of water splashes) and these enable them to detect drowning and distress incidents quicker. However, it is more often the case that drowning, and distress incidents are silent (e.g. distressed swimmers can be unable to call for help, while fellow swimmers may be unaware of nearby distress incidents; Coffman, 1991; Vittone & Pia, 2006). By removing the audio track in the current set of experiments we have created a much purer measure of visual search advantage, though it would be interesting to add audio to future studies to assess the impact of multi-sensory integration on task performance. Unfortunately, sound is likely to have little effect on localisation of a drowning target unless it was presented in stereo in combination with 360-degree visual stimuli.

The naturalistic nature of the stimuli used in the current set of experiments inevitably entail uncontrolled variation across many different dimensions. One of these is the total number of swimmers in the pool, and their behaviours, across the different clips. In some clips swimmers were more dispersed than in others, with clips containing groups of friends

creating a more cluttered search environment. Drowning targets also varied in location, with some close and some further away from the camera location. It is possible that some of these factors create an easier search environment, with targets being more salient (e.g. closer to the foreground, not surrounded by clutter) and easier to detect. While some measures were used to control for these factors, such as ensuring equal numbers of 'near' and 'far' targets in the studies, it would ultimately be difficult to control for all confounding variables in such a naturalistic task.

Although the current set of experiments have shown a consistent effect of lifeguard superiority in both accuracy of responses and in time to respond to drowning events, one limitation of the research is the ability to engage with the stimuli. When lifeguarding, surveillance is often proactive in preventing drowning in the first instance with lifeguards able to talk to swimmers or stop them entering dangerous situations (such as non-swimmers entering deep water), therefore in future research it may be interesting to explore a wider range of lifeguard behaviours (e.g. asking lifeguards 'What would you do in this situation?').

Conclusions

Despite the inevitable limitations that come with naturalistic stimuli, the research presented in these studies has highlighted that lifeguards are able to apply their knowledge in drowning detection to incidents that happen after drowning onset, suggesting the importance of lifeguard experience. This study also demonstrates an issue when there are large numbers of swimmers. Both participants groups showed degraded performance in the high set size, with lifeguard superiority attenuated. Eye-tracking measures also did not show any clear difference between lifeguards and non-lifeguards. This suggests that lifeguards'

superior drowning detection compared to non-lifeguards cannot be easily attributed to a better scanning strategy, though faster processing of drowning behaviours remains a strong possibility. Certainly, the verification time of lifeguards when fixated on targets produced the largest difference, though this failed to reach the threshold of statistical significance. However, a combination of small gains in the verification time and the time to first fixate the target potentially add up to reflect the significant response time advantage demonstrated by the lifeguard participants (echoing Laxton et al., submitted). It was also noted that the occlusion method of testing lifeguard visual search produced a greater differentiation (in terms of effect size) between lifeguard and non-lifeguard performance, suggesting that this may be a more robust way of testing lifeguard drowning detection expertise in the future. This research could be used to form the basis of an assessment tool to support lifeguard selection and training and, with further development, could provide tools for the creation of nationally consistent lifeguard training in drowning detection.

REFERENCES

Alexander, R. G., & Zelinsky, G. J. (2012). Effects of part-based similarity on visual search: The Frankenbear experiment. *Vision Research, 54,* 20-30. DOI:10.1016/j.visres.2011.12.004.

Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PloS One, 8(6),* e66169.

Biggs, A. T., & Mitroff, S. R. (2014). Improving the efficacy of security screening tasks: A review of visual search challenges and ways to mitigate their adverse effects. *Applied Cognitive Psychology*, DOI: 10.1002/acp.3083.

Borowsky, A., & Oron-Gilad, T. (2013). Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. *Accident Analysis & Prevention, 59,* 548-565. DOI: 10.1016/j.aap.2013.07.008.

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 699.

Brockmole, J. R., & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology*, *59*(7), 1177-1187.

Castro, C., Padilla, J. L., Roca, J., Benítez, I., García-Fernández, P., Estévez, B., & Crundall, D. (2014). Development and validation of the Spanish hazard perception test. *Traffic Injury Prevention*, *15*(8), 817-826.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170-178.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28-71.

Coffman, S. P. (1991). Parent education for drowning prevention. *Journal of Pediatric Health Care*, *5*(3), 141-146.

Crundall, D., Chapman, P., Phelps, N., & Underwood, G. (2003). Eye movements and hazard perception in police pursuit and emergency response driving. *Journal of Experimental Psychology: Applied*, *9*(3), 163.

Crundall, D. (2016). Hazard prediction discriminates between novice and experienced drivers. *Accident Analysis and Prevention, 86,* 47-58.

Crundall, D., & Eyre-Jackson, L. (2017). Predicting criminal incidents on the basis of non-verbal behaviour: The role of experience. *Security Journal*, *30*(3), 703-716.

Crundall, D., & Kroll, V. (2018). Prediction and perception of hazards in professional drivers: Does hazard perception skill differ between safe and less-safe fire-appliance drivers? *Accident Analysis & Prevention*, *121*, 335-346.

Curran, T., Gibson, L., Horne, J. H., Young, B., & Bozell, A. P. (2009). Expert image analysts show enhanced visual processing in change detection. *Psychonomic Bulletin & Review, 16(2),* 390-397.

Donnelly, N., Muhl-Richardson, A., Godwin, H. J., & Cave, K. R. (2019). Using eye movements to understand how security screeners search for threats in x-ray baggage. *Vision, 3(2),* 24.

Dosher, B. A., Han, S., & Lu, Z. L. (2010). Information-limited parallel processing in difficult heterogeneous covert visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1128.

Doyle, B., & Webber, J. (2007, September). *Sentinel- A systematic approach to the early recognition of drowning; the right response, to the right victim, at the right time*. Paper presented at World Water Safety Conference and Exhibition, Porto, Portugal. DOI: 10.13140/RG.2.1.1485.0081.

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision, 11(5),* 1-36. DOI:10.3758/PBR.17.5.710.

Endsley, M. R. (2017). Direct measurement of situation awareness: Validity and use of SAGAT. In *Situational Awareness* (pp. 129-156). Routledge.

Evans, K., Cohen, M. A., Tambouret, R., Horowitz, T., Kreindel, E., & Wolfe, J. (2011). Does visual expertise improve visual recognition memory? *Attention, Perception & Psychophysics, 73,* 30-35.

Faubert, J. (2013). Professional athletes have extraordinary skills for rapidly learning complex and neutral dynamic visual scenes. *Scientific Reports, 3(1154).* doi:10.1038/srep01154.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23(4),* 523-552. DOI 10.1007/s10648-011-9174-7.

Guest, D., & Lamberts, K. (2011). The time course of similarity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1667.

Griffiths, T, J. (2002). The vigilant lifeguard. *Aquatics International, 14(4),* 18-21.

Hess, A. S., Wismer, A. J., Bohil, C. J., & Neider, M. B. (2016). On the hunt: Searching for poorly defined camouflaged targets. *PloS ONE, 11(3)*. DOI:10.1371/journal.pone.0152502.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394,* 575–577.

Hout M. C., & Goldinger S. D. (2015). Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception & Psychophysics, 77 (1),* 128–49. DOI 10.3758/s13414-014-0764-6.

Howard, C. J., Troscianko, T., & Gilchrist, I. D. (2010). Eye-response lags during a continuous monitoring task. *Psychonomic Bulletin & Review, 17(5),* 710– 717. DOI:10.3758/pbr.17.5.710.

Howard, C. J., Troscianko, T., Gilchrist, I. D., Behera, A., & Hogg, D. C. (2013). Suspiciousness perception in dynamic scenes: A comparison of CCTV operators and novices. *Frontiers in Human Neuroscience, 7(441).* DOI: 10.3389/fnhum.2013.00441.

Jackson, L., Chapman,P., & Crundall, D. (2009). What happens next? Predicting other road users' behaviour as a functionof driving experience and processing time. *Ergonomics, 52(2),* 154-164. DOI: 10.1080/00140130802030714.

Jardine, N. L., & Moore, C. M. (2015). Losing the trees for the forest in dynamic visual search. *Journal of Experimental Psychology Human Perception & Performance, 42(5),* 617-630. DOI:10.1037/xhp0000186.

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis & Prevention, 42(3),* 827-834. DOI: 10.1016/j.aap.2009.09.022.

Kunar, M. A, & Watson, D. G. (2011). Visual search in a multi-element asynchronous dynamic (MAD) world. *Journal of Experimental Psychology. Human Perception and Performance, 37(4),* 1017–1031. http://doi.org/10.1037/a0023093.

Lanagan-Leitzel, L. K., Skow, E., & Moore, C. M. (2015). Great Expectations: Perceptual Challenges of Visual Surveillance in Lifeguarding. *Applied Cognitive Psychology, 29(3),* 425-435. DOI: 10.1002/acp.3121.

Laxton, V., & Crundall, D. (2018). The effect of lifeguard experience upon the detection of drowning victims in a realistic dynamic visual search task. *Applied Cognitive Psychology*. DOI: 10.1002/acp.3374.

Laxton, V., Crundall, D., Guest, D., & Howard, C. J. (2019). Visual search for drowning swimmers: investigating the impact of lifeguarding experience. Manuscript submitted for publication.

Lim, P. C., Sheppard, E., & Crundall, D. (2014). A predictive hazard perception paradigm differentiates driving experience cross-culturally. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 210-217.

Luck, S. J., & Hillyard, S. A. (1990). Electrophysiological evidence for parallel and serial processing during visual search. *Perception & Psychophysics, 48,* 603–617.

Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 887–904.

Mann, D. T., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology, 29(4),* 457-478.

McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1517.

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual skills in airport-security screening. *Psychological Science, 15(5),* 302-306.

Mitroff, S. R., & Biggs, A. T. (2014). The ultra-rare-item effect: Visual search for exceedingly rare items is highly susceptible to error. *Psychological Science*, *25*(1), 284-289.

Muhl-Richardson, A., Godwin, H. J., Garner, M., Hadwin, J. A., Liversedge, S. P., & Donnelly, N. (2018). Individual difference in search and monitoring for color targets in dynamic visual displays. *Journal of Experimental Psychology: Applied.* DOI: 10.1037/xap0000155.

Oken, B. S., Salinsky, M. C., & Elsas, S. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology, 117(9),* 1885-1901.

Page, J., Bates, V., Long, G., Dawes, P., & Tipton, M. (2011). Beach lifeguards: visual search patterns, detection rates and the influence of experience. *Ophthalmic Physiological Optics, 31,* 216–224. DOI: 10.1111/j.1475-1313.2011.00824.x

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40,* 1227–1268.

Peelan, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences, 18(5),* 242-250.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12(4),* 287–292.

Pia, F. (1974). Observations on the drowning of non-swimmers. *Journal of Physical Education. Warsaw,* IN: The YMCA Society of North America.

Pradhan, A.K., Crundall, D., 2017. Hazard avoidance in young novice drivers: definitions and a framework. In: Fisher, D.L., Caird, J., Horrey, W., Trick, L. (Eds.), *Handbook of Teen and Novice Drivers: Research, Practice, Policy, and Directions.* CRC Press, Boca Raton, FL.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science,12(1),* 48-55.

Savelsbergh, G. J., Williams, A. M., Kamp, J. V. D., & Ward, P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. *Journal of Sports Sciences, 20(3),* 279-287.

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology, 62(10),* 1904-1914.

Snodgrass, J. G., & Townsend, J. T. (1980). Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 330–354.

Stainer, M. J., Scott-Brown, K, C., & Tatler, B. W., (2013). Looking for trouble: a description of oculomotor search strategies during live CCTV operation. *Frontiers in Human Neuroscience, 7,* 1-9. https://doi.org/10.3389/fnhum.2013.00615.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113(4),* 766-786. DOI: 10.1037/0033-295X.113.4.766.

Ventsislavova, P., Crundall, D., Baguley, T., Castro, C., Gugliotta, A., Garcia-Fernandez, P., & Li, Q. (2019). A comparison of hazard perception and hazard prediction tests across China, Spain and the UK. *Accident Analysis & Prevention*, *122*, 268-286.

Vittone, M., & Pia, F. A. (2006). " It doesn't look like they're drowning". *On scene*, 14-14.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors, 50(3),* 433-441.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67-85.

Wolfe, J. M., Soce, A. A., & Schill, H. M. (2017). How did I miss that? Developing mixed hybrid visual search as a 'model system'for incidental finding errors in radiology. *Cognitive Research: Principles and Implications*, *2*(1), 35.DOI: 10.1186/s41235-017-0072-5.

Wu, C. C., & Wolfe, J. M. (2016). Multiple event monitoring. *Cognitive Research: Principles and Implications, 1(21),* 49-55. DOI 10.1186/s41235-016-0022-7.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *11*5*(4),* 787.