AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2021, Vol. 21, No. 8, 1691–1698
https://doi.org/10.1037/emo0001003

You Do It to Yourself: Attentional Capture by Threat-Signaling Stimuli Persists Even When Entirely Counterproductive

Samantha Mikhael[1], Poppy Watson[1], Brian A. Anderson[2], and Mike E. Le Pelley[1]
[1] School of Psychology, UNSW Sydney
[2] Department of Psychology, Texas A&M University

Recent research has demonstrated a counterproductive attentional bias toward threat-related stimuli: under conditions in which fixating on a color distractor stimulus sometimes resulted in an immediate shock, participants were nevertheless more likely to look at this threat-related distractor than a neutral distractor matched for physical salience. However, participants in that prior research may not have realized that their own actions caused delivery of aversive outcomes, such that monitoring for the threat-related distractor may not have been counterproductive from participants' perspective. In Experiment 1 of the current study, we demonstrate that the attentional bias to the threat-related distractor persists (and indeed, becomes stronger) when participants are made explicitly aware that looking at this stimulus is the sole cause of aversive events, which are otherwise avoidable. In Experiment 2 we replicate the bias in informed participants under conditions in which there is additional (reward-driven) motivation to avoid attending to distractors. Taken together with prior findings, the observation of an attentional bias toward the threat-related distractor under these explicitly counterproductive conditions provides strong support for the idea that threat-related stimuli are automatically prioritized by our attentional system.

Keywords: selective attention, attentional capture, threat, aversive conditioning

Supplemental materials: https://doi.org/10.1037/emo0001003.supp

It makes intuitive sense that our perceptual and cognitive systems should be adapted to promote rapid detection of signals of threat: early detection of imminent danger might allow us to alter our behavior so as to minimize or avoid negative consequences. Indeed, it has been hypothesized that the human attentional system has evolved so as to *automatically* prioritize processing of threat signals; thus, increasing the speed and cognitive efficiency of threat detection (Mulckhuyse, 2018).

Studies investigating this issue have typically made use of aversive conditioning: stimuli are established as signals of aversive outcomes (e.g., electric shock, white noise), and attention to these aversive conditioned stimuli (CS+) is subsequently compared with attention to otherwise similar stimuli that have not been paired with aversive outcomes (CS−). While such studies clearly demonstrate that humans do indeed prioritize signals of threat in many situations (for reviews, see Bar-Haim et al., 2007; Cisler & Koster, 2010), the question of whether such attentional biases reflect a truly automatic process remains more open. A body of research has attempted to address this issue by examining situations in which attending to threat-signaling stimuli is contrary to the demands of the task, or even has negative consequences for participants (e.g., Kim & Anderson, 2020; Mulckhuyse & Dalmaijer, 2016; Nissens et al., 2017; Schmidt et al., 2015; for a review, see Watson et al., 2019). The argument runs that any attentional bias that is observed under such conditions cannot reflect a goal-directed prioritization of threat signals (why would participants choose to attend to a stimulus when doing so is unnecessary or counterproductive?), and so must instead reflect automatic attentional capture by threat that is outside of the participant's top-down control. Recently, however, Anderson and Britton (2020) have questioned the support for automaticity provided by this body of research. They noted that during the attentional test phase of all of these studies, the CS+ still provided useful information about the likely occurrence of unavoidable aversive events. As such, participants may have continued to monitor for the presence of the CS+ in a goal-directed way, since doing so allowed them to anticipate and prepare for these events, and hence potentially reduce their aversive impact.

Anderson and Britton argued that compelling evidence for automatic prioritization required a task in which there was *no* goal-directed reason to attend to the threat-signaling stimulus—a task

This article was published Online First November 29, 2021.
Poppy Watson https://orcid.org/0000-0001-7454-413X
Mike E. Le Pelley https://orcid.org/0000-0002-5145-5502
This research was funded by Australian Research Council Grants DP170101715 and DP200101314. Poppy Watson was supported by an Australian Research Council Discovery Early Career Researcher Award (DE200100591). Samantha Mikhael and Poppy Watson contributed equally to this work. The authors declare no conflicts of interest.

All experiment code and raw data are available via the Open Science Framework at https://osf.io/ztqyv/. This study was not preregistered.

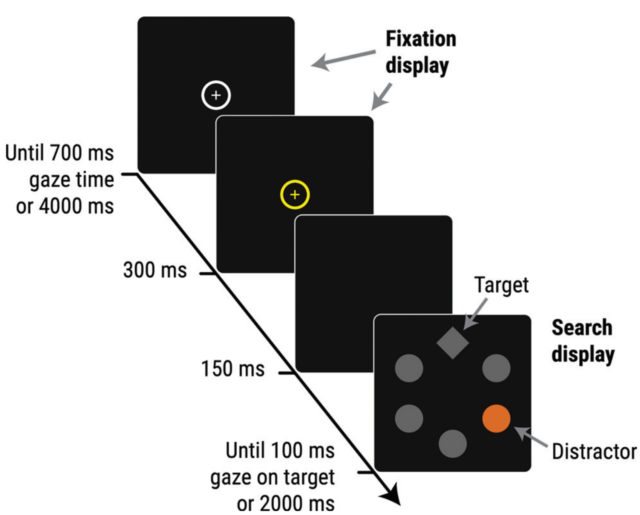
Correspondence concerning this article should be addressed to Mike E. Le Pelley, School of Psychology, UNSW Sydney, Sydney, NSW 2052, Australia. Email: m.lepelley@unsw.edu.au

in which attending to this stimulus was entirely counterproductive —and they developed a task to test this scenario. On each trial of this visual search task, participants were required to make an eye-movement (saccade) to a uniquely shaped target among a set of nontargets (see Figure 1). One of the nontarget shapes in the display could be colored either red or blue, and this colored shape was termed the *distractor*; all other shapes were gray. Critically, if participants made an eye-movement to one of the color stimuli used as distractors, there was a 50% probability that they would receive an immediate shock (CS+); gaze on the other color distractor was never paired with shock (CS−). Participants were never required to look at either distractor in this task (their task was to look at the unique shape), but looking at the CS+ was particularly counterproductive since it frequently caused shock: participants could avoid shock entirely in this task by refraining from looking at the CS+. Nevertheless, participants were more likely to look at the CS+ than the CS−. Anderson and Britton took this attentional bias—that had only negative consequences for participants—as diagnostic evidence of automatic prioritization of threat-signaling stimuli.

It is noteworthy, however, that Anderson and Britton's participants were not explicitly informed that the shock was caused by them looking at the CS+ distractor, because this raises the possibility that participants may not have realized the shock was a consequence of their own actions. Participants' gaze may have sometimes been captured by the distractor on the basis of its physical salience (as the only colored stimulus in the display: Theeuwes, 1992) and when this distractor was the CS+, this

Figure 1

Example Trial of the Visual Search Task



Note. Participants began by fixating on a central fixation cross. A search display then appeared, and participants' task was to make a saccade to a diamond-shaped target. The search display could contain a color-singleton distractor circle, colored blue or orange. Gaze on one of these colors (CS+) produced an immediate white noise burst on 50% of trials in Experiment 1 (and 100% of trials in Experiment 2). Gaze on the other color (CS–) never produced the noise. See the online article for the color version of this figure.

would sometimes be accompanied by delivery of shock. But participants may have thought this shock was inevitable, not becoming aware of the causal connection between their behavior and the occurrence of shock—especially because (a) given the rapid and dynamic nature of eye-movements, participants may not always realize that they have looked at the distractor, and (b) shock occurred on only 50% of trials in which participants looked at the CS+. As such it is possible that participants learned (erroneously) that appearance of the CS+ was a signal of occasional, unavoidable shock. Consequently it would make sense for participants to monitor for the appearance of the CS+ in a goal-directed way, because doing so would allow them to prepare for the likely shock —unaware that their own behavior increased the likelihood of shock. On this account, attentional prioritization of the CS+ in this task may still have reflected the operation of a goal-directed selection process, for similar reasons as Anderson and Britton previously highlighted with regard to earlier studies (e.g., Mulckhuyse & Dalmaijer, 2016; Nissens et al., 2017).

The issue outlined above raises the question of whether participants would still show attentional prioritization of the CS+ if they knew the full nature of the relationship between the CS+ and shock—if they knew that looking at the CS+ was the only possible cause of shock—because it is this knowledge that makes attending to the CS+ counterproductive *from the participant's perspective*. By contrast, while participants in some previous studies were explicitly informed that a failure to quickly look away from trained colors would increase the probability of shock and were motivated to look away from (rather than toward) the CS+ (Kim & Anderson, 2020; Nissens et al., 2017), the inevitability of shock on some trials of those tasks leaves open the possibility that participants were explicitly monitoring for signals of threat as described earlier. Observing an attentional bias toward the CS+ under conditions in which participants are aware of the relationship between the CS+ and aversive outcomes *and* these aversive outcomes are entirely avoidable would strengthen support for the idea that this bias reflects truly automatic prioritization of threat-related stimuli.

Experiment 1

To address this issue, Experiment 1 investigated the impact of explicit causal knowledge on attentional bias to threat signals. The first phase of the task (termed the *BeforeInfo phase*) was similar to that of Anderson and Britton, except that we used a white noise burst instead of shock as an aversive outcome. Noise bursts reliably elicit an involuntary startle response (Vrana et al., 1988) and have been used as threat-related outcomes in previous studies of aversive conditioning (Koster et al., 2004; Smith et al., 2006). After this first phase, half of the participants (*Full-Info* group) were explicitly informed that the noise was caused only by them looking at the CS+, so it could be avoided entirely if they avoided looking at the CS+. The other half of participants (*Control* group) were not explicitly informed of this causal relationship. All participants then continued with a second phase of the task (*AfterInfo phase*) as before. If any attentional bias toward the CS+ in the BeforeInfo phase reflected goal-directed selection (due to participants' misconstrual of the noise as unavoidable), then providing explicit knowledge of the true causal relationship should result in attenuation of this bias in the AfterInfo phase. By contrast, if attentional bias toward the CS+ reflected automatic prioritization of the threat-signal, it should not be influenced by participants' explicit knowledge of the causal relationship between CS+ and noise.

Method

Participants

We aimed to recruit 30 participants per group, to match the 30 participants tested by Anderson and Britton (2020). Power analysis using G*Power (Faul et al., 2007) indicated that this would provide power of >.95 to detect a medium-sized effect (η_p^2 = .06) for the within-subjects effect of distractor type (CS+ vs CS−), and also power of >.95 for the interaction between distractor type and information group (Full-Info vs. Control): default G*Power settings for correlation among repeated measures and nonsphericity correction were used. In total we tested 63 UNSW Sydney students, who participated for course credit. Two participants were excluded because the eye-tracker could not track their gaze, and one participant withdrew upon hearing a sample of the white noise burst. The remaining 60 participants (age M = 18.9 years, 95% confidence interval, CI [18.3, 19.6]; 43 female, 17 male) were alternately allocated to the Full-Info and Control groups (n = 30 per group). All research reported here was approved by the UNSW Sydney Human Research Ethics Advisory Panel (Psychology). This study was not preregistered.

Apparatus

Participants were tested using a Tobii TX-300 eye-tracker (sampling frequency 300 Hz), mounted on a 23-in. monitor (1920 × 1080 resolution, 60 Hz refresh rate), with a chin rest ~60 cm from the screen. For gaze-contingent calculations, gaze data were down-sampled to 100 Hz. Auditory stimuli were delivered via headphones (AKG K77 Perception). Stimulus presentation in the visual search task was controlled by MATLAB with Psychophysics Toolbox extensions (Kleiner et al., 2007), and the final questionnaires were presented using Inquisit.

Stimuli and Design

All stimuli were presented on a black background. Each trial (see Figure 1) began with a fixation display consisting of a central white cross inside a white circle (radius 1.5° visual angle). When participants had accumulated 700 ms gaze dwell time in this circle (or after 4,000 ms) the fixation display turned yellow to indicate the imminent search display. This search display comprised six shapes (2.3° × 2.3°)—five circles and one diamond (the target)—arranged evenly around screen center at 5.1° eccentricity. One of the circles could be colored either blue or orange (CIE *x*/*y* chromaticity coordinates: blue .192/.216, orange .493/.445, luminance ~24.5 cd/m^2): this colored circle was termed the *distractor*. For half of participants, blue was the CS+ color and orange was the CS– color; for remaining participants, this was reversed. All other shapes were gray (CIE coordinates .327/.400, luminance ~8.3 cd/m^2).

Each block of the visual search task contained 72 trials: 30 trials with a CS+ distractor in the search display, 30 trials with a CS– distractor, and 12 *distractor-absent* trials in which no color-singleton distractor was present (i.e., all shapes in the search display were gray). Trial order within each block was random, and the location of the target and distractor were randomly chosen on each trial.

A circular region of interest (ROI) with radius 1.75° was defined around the target in the search display, and a larger ROI (radius 2.55°) was defined around the distractor. The search display terminated when a response was registered (defined as 100 ms of gaze dwell time accumulated within the target ROI) or after 2,000 ms (timeout). On a random half of the trials featuring a CS+ distractor, a 50 ms, 95 dBA, white noise burst occurred immediately if any gaze was detected inside the distractor ROI. White noise with these properties has been categorized as "aversive high-intensity" (Donnerstein & Wilson, 1976) and has been used in prior studies of aversive conditioning (e.g., Hintze et al., 2014). By contrast, no noise bursts ever occurred on trials with a CS– distractor, or distractor-absent trials, regardless of participants' gaze.

If the search display timed-out with no response registered, the feedback "Too Slow" appeared for 2,000 ms; no feedback was presented otherwise. The next trial then began after a blank intertrial interval of 1,300 ms.

Procedure

All participants were initially played a sample of the white noise burst, and were told that they would sometimes hear this during the experiment. Instructions then introduced the visual search task: participants were told their task was to look at the diamond shape as quickly and directly as possible (no further information on the noise was provided at this point). Participants then completed the BeforeInfo phase of the task, which comprised four blocks of trials (structured as described earlier).

Following this first phase of the task, participants in the Control group saw a repeat of the instructions from the beginning of the experiment, stating they should continue to look at the diamond as quickly and directly as possible. Participants in the Full-Info group were instead explicitly informed of the consequences of looking at the colored circles. For example, a participant for whom the CS+ color was blue and CS– was orange would be told: "*You will have noticed that you sometimes hear a loud noise in the headphones. This only ever occurs if you look at the BLUE circle! If you can avoid looking at the blue circle you will avoid hearing the loud noise. So, you should try to move your eyes straight to the diamond. You will never hear any noise if you look at the ORANGE circle, but you should still try to move your eyes straight to the diamond*." Check questions were used to ensure that participants had understood these instructions. All participants then completed three blocks in the AfterInfo phase of the task.

Following the visual search task, participants were asked to rate how unpleasant they found the noise on a scale from 0 (pleasant) to 100 (very unpleasant). We also assessed participants' explicit knowledge regarding the consequences of looking at the colored distractors. First they were asked: "*What do you think caused the loud noise to play through the headphones?*," with response options: "*Nothing—the noise occurred randomly*" or "*I caused the noise*." A second question then asked: "*You may have noticed that the loud noise occurred when you looked at one of the colored circles. Do you know which one?*." Participants were then prompted to select the color (blue or orange) they thought was associated with the noise.

Data Preprocessing

Preprocessing of data from the visual search task followed our previous protocols (e.g., Le Pelley et al., 2015; Pearson et al., 2016; Watson et al., 2019). Data from the first two trials of the task and the first two trials after each break were discarded, as were trials that timed-out with no response (2.1% of all trials), and trials with less than 25% valid gaze data (0.7% of all trials). Across remaining trials, valid gaze location was registered in M = 98.2%, 95% CI [97.1, 99.2], of samples from the eye-tracker. Our primary measure was the proportion of trials in which gaze was recorded on the colored distractor circle: we term these *distractor-gaze trials*. The mean proportion of distractor-gaze trials was calculated separately in the BeforeInfo and AfterInfo phases for trials with a CS+ distractor versus a CS– distractor. For the sake of brevity, we do not report data from distractor-absent trials here since they do not bear on the central question of attentional bias toward threat-signaling stimuli. All experiment code and raw data are available via the Open Science Framework (OSF) at https://osf.io/ztqyv/.

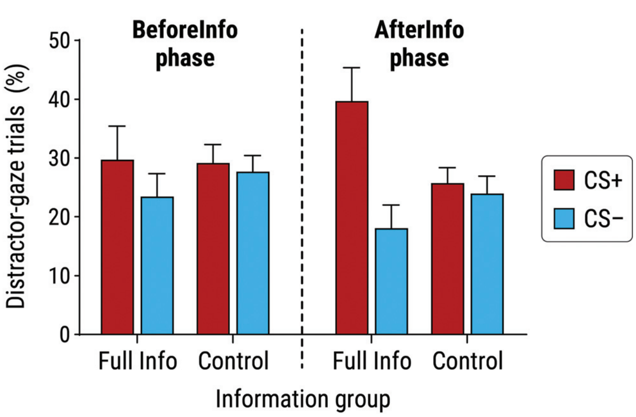
Results

For consistency with Anderson and Britton (2020) study (on which the current study was based)—as well as our own previous work using related procedures (e.g., Le Pelley et al., 2015, 2019; Pearson et al., 2015; Watson et al., 2019)—our primary focus was on the proportion of trials on which participants looked at the color-singleton distractor. In online supplementary materials we report results of secondary analyses of (a) saccade latencies for saccades made toward the target and the distractor, and (b) duration of gaze dwell time on the CS+ and CS– for the subset of trial on which participants looked at the distractor.

Figure 2 shows the proportion of distractor-gaze trials in the BeforeInfo and AfterInfo phases, for participants in the Full-Info and Control groups. These data were initially analyzed using analysis of variance (ANOVA) with a between-subjects factor of information group (Full-Info vs Control), and within-subjects factors of distractor (CS+ vs CS–) and phase (BeforeInfo vs AfterInfo). This analysis revealed a significant main effect of distractor, $F(1, 58) = 21.6$, $p < .001$, $\eta_p^2 = .272$, with more distractor-gaze trials when the display contained a CS+ than a CS– (suggesting an overall attentional bias to the threat signal). Notably, however, this effect was moderated by a significant three-way interaction, $F(1, 58) = 11.9$, $p = .001$, $\eta_p^2 = .170$. To decompose this interaction, we analyzed the data from each phase using separate Information Group × Distractor ANOVAs.

Figure 2

Proportion of Trials in Experiment 1 on Which Participants Looked at the Color-Singleton Distractor During the Search Task



Note. Data are shown separately for each distractor condition (CS+ and CS−), and for the BeforeInfo phase (before any instructions about the relationship between the CS+ and noise) and the AfterInfo phase (after the Full-Info group had been instructed that looking at the CS+ caused the noise; the Control group were never informed of this relationship). Error bars show within-subjects 95% confidence interval (Morey, 2008). CS = conditioned stimuli. See the online article for the color version of this figure.

Analysis of the data from the BeforeInfo phase revealed that the main effect of distractor approached significance, $F(1, 58) = 3.68$, $p = .060$, $\eta_p^2 = .060$, with a trend toward more distractor-gaze trials when the display contained a CS+ than a CS– (suggesting an attentional bias to the threat signal). There was no significant main effect of information group, $F(1, 58) = .39$, $p = .535$, $\eta_p^2 = .007$, or Information Group × Distractor interaction, $F(1, 58) = 1.39$, $p = .244$, $\eta_p^2 = .023$, which is unsurprising given that both groups were treated equivalently until after this first phase of the task.

Analysis of the data from the AfterInfo phase again found no main effect of information group, $F(1, 58) = 1.60$, $p = .212$, $\eta_p^2 = .027$. There was a significant main effect of distractor, $F(1, 58) = 35.1$, $p < .001$, $\eta_p^2 = .377$, and critically this was moderated by a significant interaction with information group, $F(1, 58) = 25.2$, $p < .001$, $\eta_p^2 = .303$. Figure 2 shows that, during the AfterInfo phase, the attentional bias to the threat signal (given by the difference in performance for CS+ and CS– trials) was significantly greater for the Full-Info group than the Control group.

We also analyzed data from the questionnaires administered after the visual search task. Both groups rated the noise as being highly unpleasant on the 100-point scale where 100 represented *very unpleasant* (Full-Info: M = 78.9, 95% CI [73.6, 84.2]; Control: M = 71.0, 95% CI [62.9, 79.0]), with no significant difference between the groups, $t(58) = 1.70$, $p = .095$, $d = .44$. When asked whether the noise had occurred randomly or been caused by their own behavior, only 20% of participants in the Control group correctly indicated that they had caused the noise, versus 83% in the Full-Info group (who had been explicitly told that this was the case), $\chi^2(1) = 24.1$, $p < .001$. By contrast, participants in both groups were generally able to identify which distractor color had been paired with the noise (Control: 80% correct; Full-Info: 100% correct; chance performance = 50% correct), though once again performance was significantly better in the Full-Info group who had been explicitly told the color–noise relationship, $\chi^2(1) = 4.61$, $p = .031$ (Yates correction applied).

Discussion

During the BeforeInfo phase of Experiment 1, we observed some evidence for an attentional bias to the threat-signaling distractor: there was a trend toward participants being more likely to look at the CS+ than the CS– under conditions in which looking at the CS+ had a 50% chance of resulting in an aversive noise. This finding approached significance (p = .060 two-tailed, d = .25), though we note that as a conceptual replication of Anderson and Britton (2020) the direction of the effect (CS+ > CS–) is anticipated by that prior study, and hence a one-tailed test could be justified here (that would yield a significant difference, $p = .030$). Regardless, the effect size of the attentional bias was somewhat weaker than that previously reported by Anderson and Britton ($d = .48$). We note that participants experienced more CS+ trials in Anderson and Britton's study than in the BeforeInfo phase of the current task (180 vs. 120); assuming the bias takes some amount of training to emerge, the longer task used by Anderson and Britton may have contributed to the larger attentional bias to the CS+ seen in their study. That said, for the Control group there was no difference between the BeforeInfo and AfterInfo phases of the current task, and when restricted to this group a combined analysis of these two phases (that included 210 CS+ trials, i.e., more than in Anderson and Britton's study) also did not find significant evidence of an attentional bias to the CS+ versus the CS–, $t(29) = .80$, $p = .43$, $d_z = .15$. Another possibility is that the difference in the effect size in our study versus that of Anderson and Britton may be a consequence of the difference in aversive outcome: the current study used a loud noise, whereas Anderson and Britton used electric shock, which may have been more aversive for participants and hence likely to promote a larger attentional bias. This possibility could be assessed in future research.

The more important finding of the current study related to the effect on this attentional bias of participants' explicit knowledge that it was their own behavior that caused the noise to occur. Two findings stand out here. First, very few participants in the Control group (who received instructions similar to those used by Anderson and Britton) reported being aware that they had been causing the noise, even though the noise was entirely contingent on their behavior and (when scheduled) was delivered immediately when gaze fell on the CS+ distractor. Thus, despite considerable experience of the noise ($M = 29.6$ noise events per participant, 95% CI [25.1, 34.1]), in the absence of explicit instruction participants typically did not realize that they were the cause of the noise—even though the majority (80%) of these Control participants could identify the noise-paired color when prompted. The implication is that many participants in the Control group may not have realized that monitoring for the noise-signaling CS+ was a counterproductive strategy in this task.

The second critical finding relates to the effect of instructing participants in the Full-Info group that looking at the CS+ caused the noise to occur. Contrary to our hypotheses, the effect of this instruction was to *increase* the size of the attentional bias toward the CS+. Even though participants in this group knew that they could avoid the noise entirely by not looking at the CS+, and reported finding the noise aversive, they became considerably more likely to look at the CS+ than the "safe" CS– during the AfterInfo phase.

One possibility is that the strong bias in the Full-Info group may have reflected a strategic search for information regarding the true CS+/noise relationship. Although participants in this group were explicitly told that the noise would occur only if they looked at the CS+, 17% of them still reported at the end of the experiment that the noise had occurred randomly, rather than reporting (as they had been informed) that they caused the noise to occur. This raises the possibility that (some) participants in the Full-Info group may have been confused by the 50% partial reinforcement schedule, because on half of CS+ trials, looking at this distractor would *not* cause the noise. Participants may have deliberately continued to look at the CS+ in an attempt to work out exactly when the noise would occur (an impossible task, because noise delivery was stochastic). That is, participants' goal of understanding the true nature of the CS+/noise relationship may have outweighed their goal of avoiding the unpleasant noise. By contrast, the instruction that the noise would occur *only* if participants looked at the CS+ makes the status of the CS– clear and unambiguous—the noise would never occur when the CS– was present in the display—and so there was presumably little drive for participants to explore this relationship further. As such, any "exploration-motivated" behavior would tend to favor gaze on the CS+ versus the CS– following instructions, potentially explaining the pattern of findings observed in the Full-Info group of Experiment 1.

To address this issue, Experiment 2 used full reinforcement such that looking at the CS+ *always* resulted in immediate delivery of the noise. This manipulation should have eliminated any uncertainty regarding the cause of the noise in informed participants. Furthermore, in Experiment 2 we provided participants with additional motivation to ignore distractors, by providing a monetary reward for a rapid saccade to the target on each trial. Hence there were now two reasons for participants to try to avoid looking at the CS+ distractor: because doing so always caused an unpleasant noise, and resulted in a slower response to the target and hence a lower likelihood of receiving reward. If we were still to observe an attentional bias to the CS+ under these conditions in which such a bias was even more counterproductive, this would provide stronger evidence for an involvement of involuntary processes.

Experiment 2

Method

Participants and Apparatus

Thirty UNSW Sydney participants were recruited for Experiment 2, for course credit. Two participants could not complete the study as the eye-tracker was unable to track their gaze, and data from a further participant were excluded as noises occurred at a reduced sound level due to experimenter error. The final sample comprised 27 participants (age $M = 18.9$ years, 95% CI [18.4, 19.3]; 21 female, 6 male). Apparatus was as for Experiment 1. Participants who completed the task received a monetary reward based on their performance in the rewarded phase of the visual search task ($M = 7.18$ AUD, 95% CI [6.88, 7.48]).

Stimuli, Design, and Procedure

Stimuli, design, and procedure were as for Experiment 1, with exceptions as outlined here. All participants were informed at the outset of the experiment that looking at the CS+ distractor would cause the noise, using the same instructions as given to the Full-Info group of Experiment 1. Participants then completed four blocks (288 trials) in the *unrewarded phase* of the search task, where trials were as for Experiment 1 with the sole difference that now looking at the CS+ always resulted in immediate delivery of the white noise burst.

After completing this initial phase, participants were told that in the subsequent phase of the task they could earn points that would later be converted into a cash bonus. Instructions noted that the faster participants moved their eyes to the diamond target on each trial, the more points (and hence more money) they would earn.

Participants were also told that they would still hear the noise if they looked at the CS+ distractor during this phase.

Following these instructions, participants completed four blocks in the *reward phase* of the task. In this phase, participants earned 0.1 points for every millisecond that their response time was below 1,000 ms (e.g., a response time of 600 ms would earn 40 points). A feedback display appeared immediately after participants had made their response that showed how many points had been earned on that trial (or "Too slow" if response time was greater than 1,000 ms), and the total points earned so far in the experiment. Feedback was displayed for 1,500 ms.

After the reward phase of the visual search task, participants completed questionnaires assessing the aversiveness of the noise, and their knowledge of the relationship between distractors and noise, as in Experiment 1 (note that all participants had been explicitly informed of the distractor–noise relationship in Experiment 2, so these latter questions assess retention and understanding of task instructions).

Data Preprocessing

Gaze data were processed as for Experiment 1. Data were removed for trials that timed-out with no response (2.5% of all trials), and trials with less than 25% valid gaze data (0.9% of all trials). Across remaining trials, valid gaze location was registered in M = 96.8%, 95% CI [95.5, 98.2], of samples from the eye-tracker. All experiment code and raw data are available via the OSF at https://osf.io/ztqyv/.

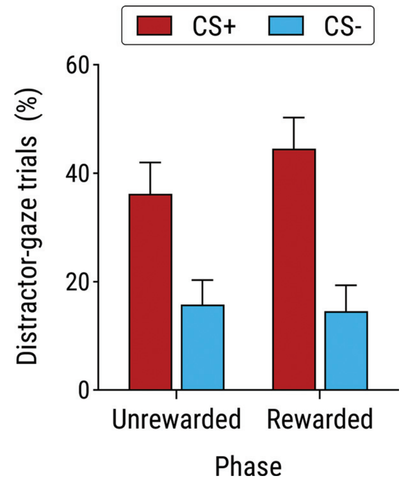
Results

As for Experiment 1, the analyses reported here focus on the proportion of distractor-gaze trials. In online supplementary materials we report results of secondary analyses of saccade latencies and gaze dwell times.

Figure 3 shows the proportion of distractor-gaze trials in the unrewarded and rewarded phases of the visual search task. These data were analyzed using ANOVA with factors of distractor (CS+ vs CS–) and phase (unrewarded vs rewarded). This revealed a significant main effect of distractor, $F(1, 26) = 50.6$, $p < .001$, $\eta_p^2 = .661$, with a greater proportion of distractor-gaze trials when the display contained a CS+ versus a CS–, demonstrating an attentional bias to the threat signal. Phase did not exert a significant main effect, $F(1, 26) = 2.67$, $p = .114$, $\eta_p^2 = .093$. However, there was a Significant Distractor × Phase interaction, $F(1, 26) = 10.2$, $p = .004$, $\eta_p^2 = .282$, with a greater attentional bias to the CS+ in the rewarded phase than the unrewarded phase. Post hoc analyses using Tukey's honest significant difference (HSD) test revealed that the effect of distractor was significant in both the unrewarded phase, $t(34.9) = 5.32$, $p < .001$, and the rewarded phase, $t(34.9) = 7.80$, $p < .001$. For the CS+, distractor-gaze trials increased in the rewarded phase relative to unrewarded phase, $t(46.1) = 3.16$, $p = .0143$; for the CS–, there was no significant difference in performance between phases, $t(46.1) = .47$, $p = .967$. The pattern of greater distraction by the CS+ than the CS– came at a cost to participants during the rewarded phase: response time to look at the target was significantly greater on trials with a CS+ (M = 477 ms, 95% CI [457, 497]) than with a CS– (M = 409 ms, 95% CI [390, 427]), $t(26) = 8.31$, $p < .001$, $d_z = 1.60$, which meant that participants earned fewer points (and hence less money) on trials with the CS+ than the CS–, because response time determined reward.

Figure 3

Proportion of Trials in Experiment 2 on Which Participants Looked at the Color-Singleton Distractor During the Search Task



Note. Data are shown separately, for each distractor condition (CS+ and CS–), in the unrewarded and rewarded phase of the visual search task. Error bars show within-subjects 95% confidence interval (Morey, 2008). CS = conditioned stimuli. See the online article for the color version of this figure.

As in Experiment 1, in postexperiment questionnaires participants rated the noise as being highly aversive (M = 84.7, 95% CI [79.5, 89.9], on the 100-point scale). All participants correctly responded that the noise had been caused by their behavior (rather than occurring randomly), and all correctly identified the CS+ color.

Discussion

In Experiment 2, all participants were explicitly informed at the outset that the noise would occur only if they looked at the CS+ distractor, and this instrumental relationship was 100% consistent: on every CS+ trial, if the participant looked at the colored distractor, the noise occurred immediately. This should have eliminated any uncertainty about the CS+/noise relationship, removing any drive to explore the boundaries of this relationship (compare with the 50% reinforcement schedule used in Experiment 1, which leaves residual uncertainty). Nevertheless, Experiment 2 again found a significant attentional bias toward the CS+ distractor versus the CS–, indicating that attentional bias toward the threat-related distractor in this task is not (purely) a consequence of strategic monitoring in an attempt to resolve uncertainty about the CS+/noise relationship. Furthermore, in Experiment 2, this attentional bias persisted—and in fact became stronger—in the latter half of the task in which we introduced a monetary reward for rapid eye movements to the target. This reward should have provided even greater motivation for participants to try to avoid attending to colored distractors: not only could this produce the aversive noise, but it would also slow responses to the target and hence reduce the money they earned. But even under these "extra-counterproductive" conditions, participants continued to show an attentional bias to the threat-signaling distractor.

It is tempting to conclude from the significant Distractor × Phase interaction that the introduction of reward increased the magnitude of attentional bias to the CS+: that the prospect of reward somehow potentiated attention to a signal of threat. However, we are wary of drawing this conclusion, because the reward manipulation was confounded with the order of the phases. An alternative possibility is that the greater bias during the second (reward) phase reflects participants' greater experience of the CS+/noise relationship during this phase. Regardless, the important finding for current purposes is that the introduction of reward did not eliminate (or even seem to reduce) the counterproductive attentional bias toward the CS+.

General Discussion

Recent research has investigated the potential automaticity of attentional bias to threat-signaling stimuli by examining performance under conditions in which attending to a CS+ distractor was (from the experimenter's perspective) counterproductive, since it caused delivery of an aversive outcome (Anderson & Britton, 2020). However, participants in that prior research may not have realized that attending to the CS+ was counterproductive, raising the possibility that they may have perceived the CS+ as a useful signal that allowed them to predict and prepare for an unavoidable shock. The current experiments ruled out this interpretation, by demonstrating that the attentional bias to the aversive CS+ persisted even when it was ensured (through instruction) that participants knew the causal relationship between their behavior and the aversive outcome (here a loud noise), and hence were aware that this outcome could be avoided entirely if they managed to avoid attending to the noise-signaling (CS+) distractor. Experiment 1 showed that instruction as to the causal nature of the relationship between looking at the CS+ and delivery of noise actually resulted in a stronger bias to the CS+ than in participants who did not receive this instruction. Experiment 2 again showed a strong attentional bias to the CS+ in informed participants, even when there was no ambiguity in the behavior–threat relationship (looking at the CS+ always produced immediate noise) and participants had additional motivation to try to ignore the colored distractors (because attending to these distractors would slow responses to the target and hence reduce reward that could be earned).

Under the conditions of the current experiments, there was no goal-directed, strategic reason for participants to choose to attend to the CS+ (and to continue to do so over the course of several hundred trials): doing so had only negative consequences, and participants were aware of this. Consequently, our findings strongly support the idea that the persistent attentional bias to the CS+ observed under these conditions reflects the operation of an automatic, involuntary attentional process that acts to prioritize threat-signaling stimuli (Mulckhuyse, 2018; Vuilleumier, 2005).

A notable finding of Experiment 1 is that instruction regarding the counterproductive consequences of looking at the CS+ resulted in a significantly *stronger* attentional bias to this stimulus, relative to the control group who did not receive this instruction. The results of Experiment 2 further demonstrated that this increase in attentional bias to the CS+ following instruction was not purely driven by ambiguity or uncertainty surrounding the partial reinforcement used in Experiment 1. Instead we interpret these findings to suggest that the explicit instruction highlighted the relationship between the CS+ and the aversive noise, clarifying for instructed participants that the CS+ was the (only) signal of threat and thereby making it more difficult to ignore. While it is true that most of the participants in the uninformed Control group were also able to correctly identify the noise-paired color when prompted to do so in the postexperiment questionnaire (suggesting some level of predictive knowledge even in these participants), they may not have reflected on this relationship during the search task itself to the same extent as those in the Full-Info group, and/or presumably did not hold the knowledge of the specific relationship between CS+ and noise with the same degree of confidence. On this account, the stronger attentional bias to the CS+ in the Full-Info group reflects stronger knowledge that the CS+ was indeed a threat signal, which in turn would fuel greater automatic attentional prioritization of this signal.

In essence this view sees the automatic attentional selection of threat signals as a form of *ironic process* (an attentional "white bear" effect: Cunningham & Egeth, 2016; Moher & Egeth, 2012; Wegner, 1994): knowledge that the CS+ signals threat causes the attentional system to monitor for this stimulus, and when such a stimulus is detected, it is particularly difficult to suppress. In drawing a link to the idea of ironic processes, we should clarify that the attentional bias observed here does not simply result from an instruction not to look at the distractor: indeed, participants were instructed that they should not look at *either* distractor (CS+ or CS–) and yet they were more likely to look at the CS+. That is, the attentional bias was tied to the aversive consequences of looking at the CS+. In this regard our findings complement and extend those of Anderson and Britton (2020). That prior study demonstrated an attentional bias under conditions in which participants received no instruction about the consequences of looking at the CS+, further ruling out interpretation of the bias as being created by the explicit instruction to avoid the distractors. Instead, the current findings and those of Anderson and Britton are consistent with the idea that the attentional bias to the CS+ reflects the operation of an automatic, involuntary process that is driven by knowledge (acquired by experience and potentiated by instruction) of the relationship between the CS+ and its aversive consequences. One possibility is that this prioritization of threat-signals reflects a fundamental drive in the attentional system that is based on *information-seeking* (Gottlieb et al., 2013, 2014): the system automatically prioritizes rapid detection and processing of stimuli that allow it (or have allowed it in the past) to anticipate aversive events, even when prioritizing such stimuli comes at a cost.

This automatic prioritization of stimuli that provide information about motivationally significant events extends beyond threat. A substantial body of existing research has also demonstrated automatic attentional capture by signals of reward (for reviews, see: Anderson, 2016; Le Pelley et al., 2016; Rusz et al., 2020; see also Watson et al., 2020). Taken together, the implication is that effects of associative learning on attention are driven by the motivational salience of the predicted outcome (i.e., the extent to which that outcome will motivate behavior) rather than the valence of the outcome (whether it is appetitive or aversive: see Watson et al., 2019). The result is an attentional system that is prioritized for rapid and cognitively efficient detection of stimuli that are likely to be important for further analysis and action.

References

Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, *1369*(1), 24–39. https://doi.org/10.1111/nyas.12957

Anderson, B. A., & Britton, M. K. (2020). On the automaticity of attentional orienting to threatening stimuli. *Emotion*, *20*(6), 1109–1112. https://doi.org/10.1037/emo0000596

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, *30*(2), 203–216. https://doi.org/10.1016/j.cpr.2009.11.003

Cunningham, C. A., & Egeth, H. E. (2016). Taming the white bear: Initial costs and eventual benefits of distractor inhibition. *Psychological Science*, *27*(4), 476–485. https://doi.org/10.1177/0956797615626564

Donnerstein, E., & Wilson, D. W. (1976). Effects of noise and perceived control on ongoing and subsequent aggressive behavior. *Journal of Personality and Social Psychology*, *34*(5), 774–781. https://doi.org/10.1037/0022-3514.34.5.774

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Gottlieb, J., Hayhoe, M., Hikosaka, O., & Rangel, A. (2014). Attention, reward, and information seeking. *The Journal of Neuroscience*, *34*, 15497–15504. https://doi.org/10.1523/JNEUROSCI.3270-14.2014

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. https://doi.org/10.1016/j.tics.2013.09.001

Hintze, P., Junghöfer, M., & Bruchmann, M. (2014). Evidence for rapid prefrontal emotional evaluation from visual evoked responses to conditioned gratings. *Biological Psychology*, *99*, 125–136. https://doi.org/10.1016/j.biopsycho.2014.03.010

Kim, H., & Anderson, B. A. (2020). How does the attention system learn from aversive outcomes? *Emotion*. Advance online publication. https://doi.org/10.1037/emo0000757

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1–16. https://doi.org/10.1068/v070821

Koster, E. H. W., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*(10), 1183–1192. https://doi.org/10.1016/j.brat.2003.08.001

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, *142*(10), 1111–1140. https://doi.org/10.1037/bul0000064

Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General*, *144*(1), 158–171. https://doi.org/10.1037/xge0000037

Le Pelley, M. E., Pearson, D., Porter, A., Yee, H., & Luque, D. (2019). Oculomotor capture is influenced by expected reward value but (maybe) not predictiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *72*(2), 168–181. https://doi.org/10.1080/17470218.2017.1313874

Moher, J., & Egeth, H. E. (2012). The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Attention, Perception & Psychophysics*, *74*(8), 1590–1605. https://doi.org/10.3758/s13414-012-0358-0

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Mulckhuyse, M. (2018). The influence of emotional stimuli on the oculomotor system: A review of the literature. *Cognitive, Affective & Behavioral Neuroscience*, *18*(3), 411–425. https://doi.org/10.3758/s13415-018-0590-8

Mulckhuyse, M., & Dalmaijer, E. S. (2016). Distracted by danger: Temporal and spatial dynamics of visual selection in the presence of threat. *Cognitive, Affective & Behavioral Neuroscience*, *16*(2), 315–324. https://doi.org/10.3758/s13415-015-0391-2

Nissens, T., Failing, M., & Theeuwes, J. (2017). People look at the object they fear: Oculomotor capture by stimuli that signal threat. *Cognition and Emotion*, *31*(8), 1707–1714. https://doi.org/10.1080/02699931.2016.1248905

Pearson, D., Donkin, C., Tran, S. C., Most, S. B., & Le Pelley, M. E. (2015). Cognitive control and counterproductive oculomotor capture by reward-related stimuli. *Visual Cognition*, *23*(1–2), 41–66. https://doi.org/10.1080/13506285.2014.994252

Pearson, D., Osborn, R., Whitford, T. J., Failing, M., Theeuwes, J., & Le Pelley, M. E. (2016). Value-modulated oculomotor capture by task-irrelevant stimuli is a consequence of early competition on the saccade map. *Attention, Perception & Psychophysics*, *78*(7), 2226–2240. https://doi.org/10.3758/s13414-016-1135-2

Rusz, D., Le Pelley, M. E., Kompier, M. A. J., Mait, L., & Bijleveld, E. (2020). Reward-driven distraction: A meta-analysis. *Psychological Bulletin*, *146*(10), 872–899. https://doi.org/10.1037/bul0000296

Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition and Emotion*, *29*(4), 687–694. https://doi.org/10.1080/02699931.2014.924484

Smith, S. D., Most, S. B., Newsome, L. A., & Zald, D. H. (2006). An emotion-induced attentional blink elicited by aversively conditioned stimuli. *Emotion*, *6*(3), 523–527. https://doi.org/10.1037/1528-3542.6.3.523

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology*, *97*(4), 487–491. https://doi.org/10.1037/0021-843X.97.4.487

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Watson, P., Pearson, D., & Le Pelley, M. E. (2020). Reduced attentional capture by reward following an acute dose of alcohol. *Psychopharmacology*, *237*(12), 3625–3639. https://doi.org/10.1007/s00213-020-05641-6

Watson, P., Pearson, D., Wiers, R. W., & Le Pelley, M. E. (2019). Prioritizing pleasure and pain: Attentional capture by reward-related and punishment-related stimuli. *Current Opinion in Behavioral Sciences*, *26*, 107–113. https://doi.org/10.1016/j.cobeha.2018.12.002

Watson, P., Pearson, D., Chow, M., Theeuwes, J., Wiers, R. W., Most, S. B., & Le Pelley, M. E. (2019). Capture and control: Working memory modulates attentional capture by reward-related stimuli. *Psychological Science*, *30*(8), 1174–1185. https://doi.org/10.1177/0956797619855964

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*(1), 34–52. https://doi.org/10.1037/0033-295X.101.1.34

Received October 18, 2020
Revision received April 26, 2021
Accepted April 29, 2021 ■