

RUNNING HEAD: AFFECTIVE ATTENTIONAL SET DOES NOT INDUCE CAPTURE

**Not Looking for Any Trouble? Purely Affective Attentional Settings Do Not Induce  
Goal-Driven Attentional Capture**

Chris R.H. Brown <sup>a\*</sup>, Nick Berggren <sup>b</sup> & Sophie Forster <sup>c</sup>

<sup>a</sup> School of Psychology, University of West London, UK

<sup>b</sup> Birkbeck, University of London, UK

<sup>c</sup> School of Psychology, University of Sussex, UK

\*Corresponding author

School of Psychology, University of West London, London, Brentford, TW8 9GA, United  
Kingdom

Email: [chris.rh.brown13@gmail.com](mailto:chris.rh.brown13@gmail.com)

Word count (excluding figures, tables and abstract): 10,719

### **Abstract**

It is well established that directing goal-driven attention to a particular stimulus property (e.g. red), or a conceptual category (e.g. toys) can induce powerful involuntary capture by goal-matching stimuli. Here we tested whether broad affective search goals (e.g. for anything threat-related) could similarly induce a generalised capture to an entire matching affective category. Across four experiments, participants were instructed to search for threat-related images in a Rapid Serial Visual Presentation (RSVP) stream, while ignoring threat-related distractors presented in task-irrelevant locations. Across these experiments we found no evidence of goal-driven attentional capture by threat distractors when participants adopted a general ‘threat detection’ goal encompassing multiple sub-categories of threat (Experiments 1a, 1b). This was true even when there was partial overlap between the threat distractors and the search goal (i.e. subset of the targets matched the distractor; Experiment 2). However, when participants adopted a more specific goal for a single sub-category of threat (e.g. fearful faces), robust goal-driven capture occurred by distractors matching this sub-category (Experiment 3). These findings suggest that while affective criteria can be used in the guidance of attention, attentional settings based on affective properties alone may not induce goal-driven attentional capture. We discuss implications for recent goal-driven accounts of affective attentional biases.

**Key words:** goal-driven attention; contingent capture; threat; attentional biases; RSVP

Abstract word count: 204

**Authors' Note:** Author NB is supported by grant ES/R003459/1 from the Economic and Social Research Council (ESRC), UK.

**Significance statement:** In this investigation we extend the contingent capture literature by revealing that individuals are capable of holding an attentional set for specific affective categories, leading to attentional capture by irrelevant distractors matching this specific affective set. However, this automatic contingent capture only extends to the specific category (e.g. fearful faces). Looking out more generally for anything relevant to danger does not appear to lead to this form of automatic distraction, despite the common affective content between the attentional set and distractors. We conclude that an attentional set based purely on affective content is incapable of producing contingent capture.

**Conflict of Interests:** The authors declare that there were no conflicts of interest to report.

In daily life we are presented with an overwhelming amount of information, of which only a small proportion is selected for further processing. Over many decades of selective attention research, a rich literature has characterised this selection process as reflecting an interplay between top-down goals (i.e. what a person wishes to attend to) and stimulus-driven factors (e.g. the sensory properties of a particular object, such as brightness). Conventionally, stimulus-driven factors have been considered to underlie involuntary attentional capture effects, while goal-driven attention has often, historically, been conceptualised as a completely voluntary mechanism (e.g. Theeuwes, 2010; 2018). However, converging lines of evidence increasingly imply that goal-driven mechanisms (i.e. those initiated in response to a goal) may also play an important, if seemingly paradoxical, role in involuntary attention (Folk, Remington & Johnston, 1992).

First, it is now well-established that asking participants to adopt particular voluntary goals can lead to powerful involuntary attentional capture by goal-matching stimuli – a phenomenon often referred to as *contingent capture* (e.g. Folk et al., 1992; Folk, Remington & Wright, 1994; Folk, Leber & Egeth, 2002; Gibson & Kelsey, 1998; Leber & Egeth, 2006). Contingent capture effects do not appear limited to simple features (e.g., colour) - similar capture effects have been observed when search goals are defined by a conceptual category. For example, instructing participants to monitor a centrally presented rapid serial visual presentation (RSVP) stream for images drawn from a particular category (e.g. toys) resulted in attentional capture by category-matching distractors (e.g. rocking horse, Lego blocks, teddy bear, Slinky), even when these appeared in a peripheral task-irrelevant location (Wyble, Folk & Potter, 2013).

Contingent capture has often been discussed with respect to its important implications for experimental situations (e.g. the possibility that seemingly stimulus-driven effects could alternatively result from goal-driven strategies, Folk et al., 1992; Bacon & Egeth, 1994;

Gibson & Kelsey, 1998). However, these large and consistent goal-driven effects on involuntary attention are presumably also influential outside the laboratory. While attentional goals are typically imposed by the experimenter in laboratory settings, outside the laboratory goals may be freely chosen on the basis of motivational factors such as current concerns, mood, and desired outcomes (Dijkterhuis & Aarts, 2010; Klinger, 2013). As such, an interesting question is whether involuntary attentional capture is commonly inadvertently activated for motivationally relevant stimuli (i.e. categories of stimuli that one might commonly adopt an attentional goal for). For instance, if we generally wish to be alert to potential threats in order to avoid putting ourselves in danger, does this lead to involuntary capture by all threat-related stimuli (including those that we may wish to ignore, e.g. violent newspaper headlines)?

Consistent with this suggestion, the ability of threat-related stimuli to – at least in certain situations - involuntarily capture attention has been widely studied (see Bar-Haim et al., 2007; Cisler, Bacon & Williams, 2009; Carretie, 2014, for meta-analysis and reviews). Interestingly, this attentional bias appears to be most pronounced in individuals with high levels of anxiety - indeed Bar-Haim and colleagues' large-scale meta-analysis across 172 studies found that the effect was absent in low anxious individuals who might arguably be less motivated to attend to potential threats. Moreover, a growing line of research within the attention and emotion literature has highlighted that the rapid allocation of attention to threat related stimuli, across individuals, is not unconditional as was often posited in early models. Traditionally, the automatic prioritisation of threat was viewed as reflecting a hard-wired and inflexible stimulus-driven mechanism (e.g. Öhman, 1992). However, this view has been challenged by recent evidence that the current relevance of threat-related stimuli to a task moderates the attentional capture by task-irrelevant affective stimuli (Everaert, Spruyt & De Houwer, 2013; Vogt, De Houwer, Crombez & Van Damme, 2013; Lichtenstein-Vidne, Henik

& Safadi, 2012). Specifically, attention to threat has been found to be heightened when the threat is relevant to the performance of a task, either because it appears in a target location, participants have to explicitly respond to the threat-related stimulus, or it resembles a threat-related target. Conversely, when the threat-related stimulus is unrelated to the performance of a task, attention to these stimuli is often attenuated or abolished completely (although see Grimshaw, Kranz, Carmel, Moody & Devue, 2018).

For instance, Lichtenstein-Vidne et al. (2012) found that threat-related distractors captured attention only when the target was also threat-related and not when it was neutral (see also, Vromen, Lipp, Remington & Becker, 2016). Similarly, when performing a task with both affective (e.g. fearful faces) and neutral targets (e.g. buildings), the processing of the affective target only interferes with identification of a subsequent neutral target when participants judge the affective content of the first stimulus; this interference with subsequent responding is absent if participants have to judge the same images on non-affective features such as gender (e.g. Stein, Zwickel, Ritter, Kitzmantel & Schneider, 2009). Based on such evidence, Feature-Specific attention allocation theory (cf. Everaert et al., 2013) has suggested that attention to emotional stimuli only influences cognition when the emotional features are task-relevant, and that the affective content of a stimulus captures attention because participants are more likely to find it relevant to their aims. Manipulations of task-relevance could be viewed as indirect manipulations of top-down search goals – for example, if participants notice that targets are always threat-related, they may be more likely to adopt a goal for threat even if this was not directly instructed, potentially inducing contingent capture effects. Recently, we have built on this growing line of research on task-relevance using direct manipulations of search goals. Adapting the contingent capture paradigm of Wyble and colleagues (2013), we have demonstrated that involuntary attentional capture effects for

certain categories<sup>1</sup> of motivationally-relevant stimuli such as threatening animals, emotional faces, alcohol, and rewarding and aversive smoking cues can be experimentally induced when participants are asked to adopt a search goal for these categories. Such findings raise the possibility that the phenomenon of involuntary attentional capture by threat might conceivably reflect a goal-driven mechanism (Brown, Berggren & Forster, in press; Brown, Duka & Forster, 2018a; Brown, Forster & Duka, 2018b).

Given that it may often be adaptive to voluntarily look out for multiple potential dangers, a goal-driven account of attentional capture by threat appears plausible. For instance, when walking home late at night, we might act to protect our own safety by looking out for broken glass on the floor, a speeding taxi, or a stranger lurking in the shadows. Indeed, the idea that attentional biases towards threat may be due to an intentional vigilance has been raised in early models of anxiety (e.g. Well & Mathews, 1994). We note, however, that our previous demonstrations of goal-driven capture by motivationally relevant stimuli have been limited to single conceptual categories (e.g. threatening animals or negative smoking images; Brown et al., in press; 2018b). One question which remains, therefore, is whether or not goal-driven attentional capture can be elicited in response to a general affective goal (e.g. threat detection), and generalise across multiple sub-categories of threat (e.g. threatening objects, threat-related vehicles, and threat-related faces). On one hand, one might speculate that the general affective category of ‘threat’ is sufficiently well learned to

---

<sup>1</sup> Here we use the term 'category' simply to refer to a class of stimuli with particular shared characteristics. These shared characteristics may include perceptual features but may also be based on other properties such as conceptual or semantic associations (see Wyble et al., 2013). Following Cunningham and Wolfe (2014), we do not predict differences in visual search for superordinate versus subordinate category levels.

serve as a top-down attentional setting and induce goal-driven capture. Given that a wide range of different classes of threat stimuli produce similar patterns of attentional capture and overlapping neural activation (e.g. the amygdala; Bishop, 2007; Sander, Graffman & Zalla, 2003; Öhman, 2005), it is plausible that affective categories can be represented as a sufficiently cohesive category to act as a top-down goal. Further, there is evidence that when individuals learn an arbitrary category of abstract stimuli (i.e. novel shapes), pairing of an aversive outcome with a subset of exemplars leads to an automatic fear response (i.e. galvanic skin response, aversive outcome expectancy) to all exemplars, despite being visually dissimilar, suggesting that affective categories can exist beyond visual similarity or rigid conceptual relations (Vervoort, Vervliet, Bennett & Baeyens, 2014). Other research using non-affective stimuli indicates that well-learned but perceptually diverse categories can produce involuntary goal-driven attentional capture: Giammarco and colleagues demonstrated that after training participants to learn arbitrary categories of visually dissimilar everyday objects, instructions to search for this novel category resulted in attentional capture by task-irrelevant stimuli from this learnt category (Giammarco, Paoletti, Guild & Al-Aidroos, 2016). Thus, these two independent lines of evidence suggest that individuals are capable of learning visually heterogeneous affective categories independent of visual features, and that well-learned yet visually heterogeneous categories can elicit goal-driven attentional capture. Combined, this evidence supports the hypothesis that purely affective categories could produce involuntary goal-driven attentional capture.

On the other hand, one of our recent investigations found that although searching for threatening animal targets produced goal-driven capture by threatening animal distractors, this effect did not generalise to another form of threat - fearful faces did not capture attention during search for threatening animals (Brown et al., in press). This could suggest an important limitation to the ability of purely affective categories to capture attention, which



would have key implications for both understanding attentional capture by threat and predicting the extent of goal-driven influences on involuntary attention in daily life. On the other hand, it could be argued that our instructed search goal for the single sub-category of threatening animals was more specific than might be observed in many real-world contexts, and hence elicited a narrower form of goal-driven attentional capture. The current study, therefore, aims to provide a direct test of the possibility that adopting a general attentional goal for the category of ‘threat’ can induce attentional capture by all threat stimuli, regardless of their subordinate category.

To this end, we used the same RSVP contingent capture task as in our previous investigation measuring goal-driven capture by threat (Brown et al., in press); however, instead of instructing participants to detect a target from a specific threat category (i.e. dangerous animals), we asked them to adopt a threat-detection search goal, defined as searching for “anything which could cause or show pain, death, or signal danger”. For the comparative non-threat search goal condition, we asked participants to adopt a positive search goal, defined as searching for “anything which makes people happy or portrays positive emotion”. We predicted that when participants searched for the general category of threat-related stimuli in the central stream, peripheral or parafoveal task-irrelevant threat-related distractors, not part of the target set, would capture attention more than a neutral category of distractor. However, when participants are searching for the general positive category, we predicted that this difference should be eliminated.

### **Experiment 1a and 1b**

#### **Methods**

Experiment 1a and 1b used near identical methods with only a minor variation in the size and position of stimuli: For Experiment 1a distractor images were larger and appeared in peripheral locations of the visual field; whilst in Experiment 1b distractor images were

resized so as to appear in parafoveal locations of the visual field, in line with previous investigations which have found conceptual generalisation of contingent capture (Wyble et al., 2013).

### **Participants.**

*Experiment 1a.* 29 participants were initially recruited from the subject pool at Birkbeck University of London, though two participants were excluded prior to analysis for accuracy being 2 *SDs* below the group mean ( $M = 61.78\%$ ,  $SD = 15.05$ ), thus participants who scored below 31% accuracy were excluded. Our stopping rule was based on the maximum number of participants that could be recruited within one week. To estimate power to detect goal-driven effects, we identified five papers in which task-irrelevant categorical distractors captured attention in an RSVP task due to congruence with the search goal (Brown et al., 2018a; 2018b; in press; Giammarco et al., 2016; Wyble et al., 2013). The average effect size for the difference between target identification/ detection accuracy when a distractor category overlapped with the search goal versus when it was incongruent, was  $d_z = 1.04$  ( $SD = .52$ ;  $k = 19$ ). A power analysis using G\*power software revealed that a sample of 10 participants would be suitable to detect this average effect ( $\alpha = .05$ ,  $\beta = .80$ ; Faul, Erdfelder, Lang, & Buchner, 2007). Thus, our final sample size should be highly powered to detect distraction by conceptually overlapping distractors. The final sample consisted of 17 females and 10 males, the mean age of which was 25.48 ( $SD = 6.87$ ). Participants were remunerated with course credits or a small cash payment.

Participants across all experiments completed an informed consent procedure prior to participating in the study. All methods reported were in accordance with the declaration of Helsinki, and were approved beforehand by either the Birkbeck Psychological Sciences Ethics Committee (Experiment 1a), or the University of Sussex Science and Technology Cross-Schools Research Ethics Committee (Experiments 1b, 2 and 3).

**Experiment 1b.** To ensure sensitivity to interpret null effects, the sample size for this second experiment was determined through a Bayesian stopping rule. We therefore stopped collecting data once the Bayes factor comparing evidence for goal-driven attentional capture (i.e. the difference in accuracy between the neutral distractor and the threat distractor in the threat search condition) was either above or below the selected cut-off for a substantial evidence for either the experimental or null hypothesis. A factor above 3 shows evidence favouring the experimental hypothesis, whilst a factor below .33 shows evidence favouring the null (Dienes, 2008; 2016).<sup>2</sup>

An initial sample of 21 participants was recruited over the final few weeks of the academic term, from the subject pool at the University of Sussex. Of this initial sample, 5 participants were excluded for scoring below the 31% accuracy cut-off which we retained from Experiment 1a. The Bayesian stopping criterion was checked for the first time after the end of term and following these exclusions, which revealed that the key comparison reached sensitivity ( $B < .33$ ), thus we did not require further data to draw a conclusion. Our final sample consisted of 12 female and 4 male participants, the mean age of whom was 20.94 (*SD*

---

<sup>2</sup> A Bayesian stopping rule does not require an *a priori* sample size due to the Bayes factor being generated from the collected data. Once the Bayes factor has become sensitive (i.e.  $B < .33$ ;  $B > 3$ ) it suggests that the effect is consistent enough to provide strong evidence for either the null or the experimental hypothesis, and data collection can be halted. Variable or inconsistent data can result in non-significance despite a real effect being present in the data; a Bayes factor is sensitive to this variability (i.e. *SE* of difference) and will only favour the null or experimental hypothesis if the difference between conditions, or the absence of a difference, is sufficiently consistent. Until evidence strongly favours either hypothesis, more data must be collected (cf. Dienes, 2016; see Bayesian analysis section below).

= 3.36). The full analysis of this data (including any calculation of  $p$ -values) was conducted only when the final sample had been collected. A power analysis using G\*power software revealed that our final sample had the power of  $\beta = .80$  to detect a pairwise effect of  $d_z = .75$  or above ( $\alpha = .05$ ; Faul et al., 2007).

**Stimuli.** The stimuli were presented using E-prime 2.0 software on a 24-inch Dell monitor with a screen resolution of 1920×1080 and a refresh rate of 60 Hz in Experiment 1a, and a 16-inch Dell monitor with a screen resolution of 800×600 and refresh rate of 60 Hz in Experiment 1b. The experiment was conducted in a dimly lit room. In total there were nine positive targets and nine threat-related targets, each made up of equal numbers of objects, animals, and faces: three in each category. The images were selected to be in line with the definitions of threat-related targets (i.e. “anything which could cause or show pain, death, or signal danger”) or positive-related targets (i.e. “anything which makes people happy or portrays positive emotion”) given to the participants. The threat-related stimuli were selected based on their association with threat outcomes, as recent research has revealed that the Pavlovian association with threat is sufficient to capture attention (e.g. Schmidt, Belopolsky & Theeuwes, 2015).

The animal images were all sourced from a previous investigation by Brown et al. (in press). The threat-related animals selected for targets were a snake, a spider, and an attacking dog (this final image was sourced from the International Affective Picture System database, IAPS; Lang, Bradley & Cuthbert, 1997). The positive animals included a kitten, a puppy, and a duckling. Previous ratings from Brown et al. (in press) confirmed that the threatening animals were considered moderately to highly arousing and were rated as having a negative valence, and positive animals were moderately to highly arousing and were rated as having a positive valence.

The object images were taken from the IAPS image database or sourced from Google images. The threat-related images included a gun, a knife, and a syringe. The knife and the syringe were sourced from online due to them having better visual quality compared to the IAPS images of the same objects. The positive images were all sourced from Google images and were based on general positive categories. We sourced three images including a bunch of flowers, money in pounds, and a gift wrapped present.

The final subcategory of target stimuli were emotional faces, all these stimuli were sourced from the NimStim database which have previously been used in experiments which found attentional capture by emotional faces (Tottenham et al., 2009). The identities of the faces were the same for both happy and fearful targets, meaning that they only differed in their emotional content. These identities included two male faces and one female face. As in previous investigations which found attentional capture by emotional faces we ovalled these faces to remove non-emotional identifying features (i.e. hair and jawline).

The internal contexts (i.e. cues within the image) of potentially ambiguous stimuli were selected to favour the threat-related interpretations; all animals appeared in attack positions (e.g. bared teeth), as did objects (e.g. knife held in clenched fist). Additionally, the positive images were also selected to be unambiguously and universally positive; for instance, the objects and animals did not require extensive personal history of positive association (e.g. cigarettes in smokers).

All images taken from established image sets (i.e. IAPS) were previously rated on arousal and negative valence ratings. In line with previous investigations, threat-related images were selected if they had moderate to high arousal ratings and high negative valence ratings (e.g. Koster, Crombez, Verschuere & De Houwer, 2004; Vogt, De Houwer, Koster, Van Damme & Crombez, 2008). Stimuli not from image sets were selected based on their similarity to those in the set (see Most et al. (2005) for similar procedure).

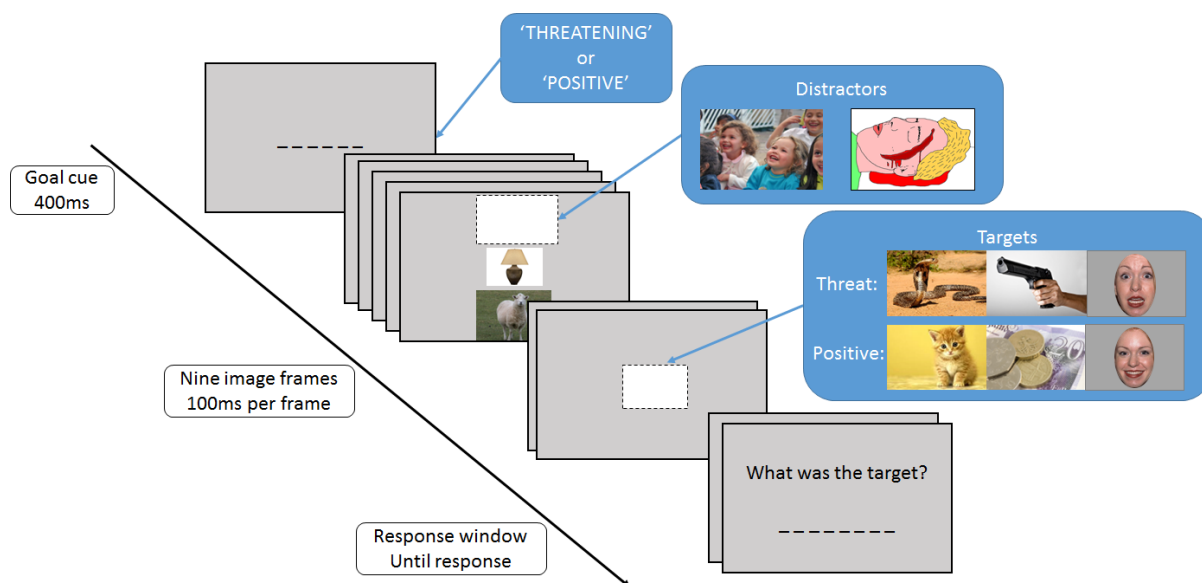
The neutral filler images presented in the task were also objects, animals, and faces. The neutral faces were an equal split of male and female faces, and were composed of a mixture of different ethnicities. In total there were 48 neutral faces which were sourced from both the NimStim image database (Tottenham et al., 2009) and the Productive Aging Laboratory Face database (Minear & Park, 2004). The neutral objects selected were everyday household items, such as shoes, bicycles, or furniture. There were a total of 48 of these neutral objects which were sourced from the IAPS image database (Lang et al., 1997). The neutral animal images were sourced from Brown et al. (in press) and were selected based on their affective ratings being neither threatening nor cute. The chosen exemplars consisted of animals such as fish, cows, pigs, and camels. All unlicensed images are listed online via the Open Science Framework ([osf.io/ju87s](https://osf.io/ju87s)). All target and filler images were presented in the centre of the screen and measured  $6^\circ \times 4.02^\circ$  in Experiment 1a and  $3.44^\circ \times 2.29^\circ$  in Experiment 1b, at a viewing distance of 59cm maintained using a chin-rest.

The distractor images measured  $8.09^\circ \times 5.35^\circ$  in Experiment 1a and  $4.58^\circ \times 2.98^\circ$  in Experiment 1b. In both experiments, distractors appeared both above and below the central RSVP stream with a gap of  $.5^\circ$  separation from the central image. The positive distractors consisted of nine images depicting people celebrating a marriage or sporting victory, or children playing. The threat-related distractors consisted of nine images of people who were fatally wounded or mutilated, such as those in a murder scene or a car accident. These affective images were all sourced from the IAPS database (Lang et al., 1997). These exact threat related images have been found to capture attention in similar RSVP tasks (e.g. Most et al., 2005; Kennedy & Most, 2015). The neutral distractor images consisted of 24 different scenes of people doing everyday activities (e.g. people shopping, at work, or on public transport). Twelve of these were sourced from the IAPS images set (Lang et al., 1997) and 12 from Google images, those taken from online were selected based on their similarity to the

neutral images taken from the IAPS database, that is, they included different images of the same content. Therefore, all distractor images depicted scenes consisting of individuals or groups, though with different affective associations.

**Procedure.** See Figure 1 for an example trial sequence in the experimental paradigm. In each block of the search task participants were given the instruction to search an RSVP stream for either a general positive category (“anything which makes people happy or portrays positive emotion”) or threat-related category (“anything which could cause or show pain, death, or signal danger”). Text search goal cues (i.e. “Positive”, “Threat”) were also presented at the beginning of each trial for 400ms. The single target stimulus was presented in a nine frame RSVP stream consisting of eight other neutral stimuli which were randomly selected from the total pool of neutral stimuli. The nine images on a particular trial were made up of three objects, three animals, and three faces, one of which was the target. This meant that when a target was from a particular category only two neutral fillers were presented from that category.

Each stimulus frame was presented for 100ms with no inter-stimulus interval. The target stimulus appeared at positions four, six, or eight in the RSVP stream an equal number of times, and was counterbalanced across conditions. The peripheral distractor stimulus was consistently presented two slides prior to the target at Lag 2. These peripheral distractors were two images presented above and below the central stimulus position. One distractor image was always a neutral distractor, randomly selected from the pool of neutral images. The other distractor stimulus could either be a threat-related distractor, positive distractor, or another neutral distractor. Within each condition the distractor image appeared an equal number of times above and below the central stream.



**Figure 1.** An example trial sequence from Experiment 1a and 1b. The first frame which appeared was a 400ms cue for the search goal for that block, and was either “THREATENING” or “POSITIVE”. This was followed by a nine frame RSVP stream, with each frame appearing for 100ms. Eight of the images were neutral, and one was the target which appeared at one of four locations (five, six, seven, or eight) in the RSVP stream, and was either a positive image (e.g. smiling children) or threat-related (e.g. mutilation or death). Distractors appeared above and below the RSVP stream and always two frames prior to the target. At the end of the trials, participants typed what they thought the positive or threat-related target was using the keyboard. In Experiment 2 and 3, the response was present/absent judgment, which participant made at the end of the trial. In the figure, the mutilation and death image is an illustration of an IAPS image created for the purpose of the figure. The emotional faces in the figure were taken from the NimStim image set with permission for publication by Tottenham and colleagues (2009).

At the end of each trial, the participant typed out the positive or threat-related image they identified as the target using the keyboard and pressed the ‘Enter’ key, which triggered the beginning of the next trial. The dependent variable was the percentage of trials where participants accurately reported the correct positive or threat related objects, faces and animals which had been presented. In total there were four blocks of 54 trials each, with a period of rest every two blocks, the duration of which was determined by the participant. The search condition blocks were presented in an alternating format (e.g. positive-threat-positive-threat). The block order was counterbalanced between participants, with half the participants



completing a threat search block first. When blocks were not separated by a rest period a text warning was presented for 3000ms alerting the participant that the search goal had changed. Other than the search goal, which was manipulated between blocks, all within participant factors were fully counterbalanced within each block.

Before the main task, participants completed a six-trial practice block, which required them to search for house images amongst a stream of cars, shoes, bricks, and trees, and type out a specific feature of each house. The specific images used in these practice trials were different from the set used in the main experiment.

**Scoring.** In order to determine the percentage of trials which were correct, the participants' responses were checked against the correct answer using an Excel formula which marked a trial as correct when the spelling of the response corresponded to the spelling of the correct answer. In order to account for spelling and approximate responses, the first author (C.R.H.B) coded the participants' responses prior to the Excel formula being applied. C.R.H.B was blind to both the distractor conditions and the correct answers during this process.

Prior to scoring, the following coding rules were applied to all responses: Incorrectly spelt answers were corrected to the most similar target included in the set of images. Vague descriptions of objects were not allowed, despite being similar to the target if it could also describe another target, e.g. "sharp object" was marked incorrect when the target was a syringe due to it also being descriptive of the knife target. Animals judged to be subordinate to the potential target animal were changed to the superordinate animal (e.g. "black widow" was accepted for spider; "cobra" was accepted for snake). The block context was taken into account, meaning that if the search condition was positive, the answer of "dog" was changed to "puppy", despite dog also being accepted as an answer in the threat search block. Due to difficulty distinguishing the different faces apart, any description of a fearful face was

accepted for all three fearful face targets, i.e. “scared”, “shocked”, “fear” were all accepted. To remain consistent across answers, changes were made universally to all answers made by a single participant, meaning that once a change was made to an answer it was also made for all identical answers that individual participant had made. The percentage of correctly identified animals was recorded as the outcome measure for analysis.

**Bayesian analysis.** To supplement the main analysis, we computed Bayes factors in order to determine whether any null effects were due to insensitivity or a true null effect. A Bayes factor compares evidence for the *experimental hypothesis* (goal-matching affective stimuli will result in greater attentional capture versus neutral stimuli) and the *null hypothesis* (goal-matching affective stimuli will not result in attentional capture versus neutral stimuli). The Bayes factor ranges from 0 to infinity, values less than 1 indicate that there is support for the null hypothesis, whilst values of greater than 1 indicate that there is support for the experimental hypothesis. The strength of this evidence is indicated by the magnitude of the Bayes Factor; values greater than three or less than .33 indicate substantial evidence for either the experimental or null hypothesis. A value closer to 1 suggests that any non-significant result is due to insensitivity and any difference or null effect is ‘anecdotal’ (Jeffreys, 1961; Dienes, 2008; 2011; 2014; 2016).

The Bayes factor was computed using a modified version of Baguley and Kaye’s (2010) R code (retrieved from Dienes, 2008). To compute the factor, we used a half-normal distribution which estimates that smaller differences are more probable than large differences. This half-normal was set with a mean of zero which reflects the null hypothesis of zero difference. we used a half-normal distribution due to the previous evidence in the literature that the effect would be directional; specifically, that threat-related stimuli would capture

attention more than neutral stimuli<sup>3</sup>. The standard deviation of this distribution was set to 8%, which reflects the plausible effect size taken from Brown et al. (in press; Experiment 1). This effect size was taken from the comparisons between the neutral baseline distractor and the threatening animal distractor condition when participants were searching for threatening animals. All direct comparisons between conditions were tested using Bayes factors, however, p-values were also computed using two-way paired samples t-tests to facilitate comparison to previous results.

### Results and Discussion

For both Experiment 1a and 1b the identification accuracy for each condition was entered in a 2×3 repeated measures ANOVA with search goal category (threat/ positive) and distractor conceptual category (threat/ positive/ neutral) as the two factors (see Table 1). We note that the accuracy reported in both experiments was in line with that reported across the four experiments in Wyble et al. (2013) using an identical response format,  $M = 53.5\%$ ,  $SD = 9.63\%$ . The main effect of search goal category was significant for both Experiment 1a and 1b, with targets from the threat-related affective search goal being more accurately identified than targets which were part of the positive affective search goal, Experiment 1a:  $F(1,26) = 21.33$ ,  $p < .001$ ,  $\eta^2_p = .45$ ; Experiment 1b:  $F(1,15) = 6.12$ ,  $p = .026$ ,  $\eta^2_p = .29$ . There was, however, no significant difference between the three distractor conceptual categories in either experiment, Experiment 1a:  $F(2,52) = .646$ ,  $p = .634$ ,  $\eta^2_p = .02$ ; Experiment 1b:  $F(2,30) = 1.09$ ,  $p = .349$ ,  $\eta^2_p = .07$ . Further, against our hypothesis, we found no significant interaction between the current search goal and the distractor conceptual category in these two

---

<sup>3</sup> Following Dienes (2008; 2011), an adjusted standard error was applied based on the sample sizes collected. This adjustment was done using the following equation:  $SE*(1 + 20/df*df)$ .

experiments, Experiment 1a:  $F(2,52) = .57, p = .568, \eta^2_p = .02$ ; Experiment 1b:  $F(2,30) = .77, p = .472, \eta^2_p = .05$ .

**Table 1.** Mean and standard deviations for percentage accuracy across all distractor and search goal conditions within Experiment 1a and 1b.

		Positive scene distractor	Mutilation scene distractor	Neutral scene distractor
Experiment 1a ( $n = 27$ )	Positive search goal	53.70 (19)	53.40 (21.29)	55.76 (19.67)
	Threat search goal	69.86 (16.50)	68.72 (16.73)	69.24 (18.07)
Experiment 1b ( $n = 16$ )	Positive search goal	53.47 (15.36)	53.99 (19.67)	55.90 (18.14)
	Threat search goal	63.89 (12.55)	65.63 (12.66)	64.76 (12.49)

In order to test the sensitivity of this null result we conducted Bayesian pairwise comparisons between the affective distractor and the neutral distractor in each condition for both experiments. This revealed that when participants were searching for the conceptual category of threat, both experiments found a sensitive null difference between threat-related and neutral distractors, Experiment 1a:  $t(26) = .30, p = .766$ , Hedges'  $g_z = .06, B_{H[0,8]} = .28$ ; Experiment 1b:  $t(15) = .60, p = .558$ , Hedges'  $g_z = -.15, B_{H[0,8]} = .13$ , thus revealing that there was evidence in favour of the null hypothesis: searching for threat as a broad conceptual category did not induce goal-driven capture by all threat-related images. There was also sensitive null differences between positive and neutral distractors in the threat search condition, Experiment 1a:  $t(26) = .29, p = .773$ , Hedges'  $g_z = -.06, B_{H[0,8]} = .21$ ; Experiment 1b:  $t(15) = .64, p = .533$ , Hedges'  $g_z = .16, B_{H[0,8]} = .31$ . A parallel analysis for the positive search condition found that the null results for this condition did not reach the threshold for a sensitive null result, but nevertheless favoured the null (Bayes factors for the equivalent contrast ranged between .65 and .84; Hedges'  $g_z$  ranged between .15 and .32). The results from Experiment 1a and 1b revealed that, despite the threat-related distractors sharing the same affective category as the broad affective search goal, participants were no more distracted by these images than the neutral images.

## Experiment 2

In the previous experiments, the threat-related distractor was never part of the target set, meaning that there was no feature overlap with the search goal, only affective overlap. Experiments 1a and 1b provide initial evidence that involuntary attentional capture cannot be caused purely by affective overlap. In these experiments, participants searched for the general category of threat, made up of three different subcategories of threat related stimuli (i.e. dangerous animals, fearful faces, and dangerous objects).

Participants did not have forewarning of the target subcategories, in order to encourage a broad search goal based purely on affective relevance. We cannot rule out, however, the possibility that participants may have become aware of the subcategories of threat and hence adopted a narrower goal for these three specific subcategories, rather than a more general goal for all threat. If participants were searching specifically for threatening animals, fearful faces, and threatening objects this could explain the lack of goal-driven capture by images of mutilation. It therefore remains possible that a broader affective search goal encompassing all possible categories of threat-related stimuli could potentially induce general goal-driven involuntary capture by threat.

Experiment 2 therefore examined whether a broad search goal encompassing multiple threat-related categories could lead to involuntary capture by stimuli from one of these categories. To test this, we included one of the target categories – fearful and neutral faces - as distractor conditions, alongside the mutilation and neutral scene distractor conditions (which as in Experiment 1 were not part of the target set). We expected that the fearful faces would interfere more with the task than the neutral faces in this Experiment because of the overlap with the search goal. Additionally, to encourage the adoption of a broad attentional set for the affective category, we increased the number of target exemplars within the threat-related target category from 9 to 24 images. We, however, expected to replicate Experiments

1a and 1b and find no difference between the mutilation and neutral scenes, even with this broader search.

In order to conserve statistical power with multiple conditions, we dropped the positive search and positive distractor condition in favour of focusing purely on threat-related categories. Another alteration to the current task was the type of response that participants made to the targets. In Experiment 1a and 1b the response required relatively deep processing of the targets, and this may have encouraged participants to adopt an overly specific goal. In Experiment 2, participants now had to make a present/ absent judgment which would have a lower threshold to correctly respond to the target (Shang & Bishop, 2000). This response format has been previously used in other variants of our task to demonstrate goal-driven capture (e.g. Brown et al., 2018a; 2018b).

## Methods

**Participants.** As in Experiment 1b, participants were recruited until all Bayes factors for the pairwise comparisons between neutral and the affective goal-congruent threat-related distractors were sensitive (i.e.  $B > 3$  or  $B < .33$ ). The Bayes factors were checked at the end of each day of testing. This led to 16 participants being recruited, of which 12 were female and 4 male. The average age of the participants was 24.31 ( $SD = 4.29$ ). Participants were remunerated with course credits or a small cash payment. A power analysis using G\*power software revealed that our final sample had the power of  $\beta = .80$  to detect a pairwise effect of  $d_z = .75$  or above ( $\alpha = .05$ ; Faul et al., 2007).

**Stimuli.** All stimuli were similar to Experiment 1a and 1b with the following exceptions. Firstly, all stimuli were now presented on a Dell 1707FP monitor with the resolution set to  $1280 \times 1024$  and a refresh rate of 60Hz. Additionally, due to the present/absent judgement task requiring more images for the increased number of trials, additional images were collected for presentation in the RSVP stream. Further, all images

within this task were taken from existing databases, in order to exclude any possibility that the previous result was influenced by the inclusion of any stimuli from unestablished sets. These images included 136 neutral animal images, 136 neutral face images, and 136 neutral objects. The neutral animals were taken from a previous investigation conducted in our lab which had been rated along dimensions of threat, cuteness, positive, and negative and were not rated highly on any of these dimensions (Brown et al., in press). The neutral faces were sourced from the Lifespan Adult Facial Stimuli Database (Minear & Park, 2004). The neutral objects were taken from the IAPS (Lang et al., 1997), Nencki Affective Picture System (NAPS; Marchewka et al., 2014) and the Geneva Affective Picture Database (GAPED; Dan-Glauser & Scherer, 2011). All image sets have been used in prior investigations of attention to emotional stimuli.

For the threat target category, eight threatening animal stimuli, eight fearful face stimuli, and eight threatening objects were presented. The threatening animals were also taken from the previous investigation and had also been rated as moderately arousing and unpleasant; the different types of animal were identical to those presented in Experiments 1a and 1b (Brown et al., in press). The fearful faces were taken from the NimStim database (Tottenham et al., 2009) and included an equal balance of male and female faces, as well as a range of ethnicities; as with the neutral faces, these were also ovalled. The majority of the threatening objects were taken from the IAPS and NAPS databases. However, to increase the number of threatening object targets, four images were sourced from Google images. These were objects which were part of the IAPS and NAPS image sets but with slightly different features (e.g. orange syringe instead of blue). The exact objects presented were a burning car, knives, syringes, and guns.

For the distractor images, twelve neutral faces and twelve fearful faces were taken from the NimStim database. These shared the same identity and retained the same balance of

genders as Experiment 1a and 1b. The mutilation and neutral scene distractor images consisted of twelve neutral scenes and twelve threat-related scenes, all taken from the IAPS image set (Lang et al., 1997). These included those presented in Experiments 1a and 1b.

As before, when the distractors were presented above or below the RSVP stream, the opposite distractor location was occupied with another image. For the face stimuli this was a patch of skin texture taken from a close-up of the distractor faces and contained no facial features. Twenty-four of these skin texture patches were presented, and included a range of different skin tones which were matched to the distractor faces. For the scene distractors the opposite distractor location was occupied with an inverted and blurred social scene sourced from Google images. Twenty-four of these images were created, and contained a similar range of colours to the neutral and threat-related scenes, but without the affective or conceptual content.

All images were the same size as in Experiment 1b, with the exception of the fearful and neutral face distractors presented in parafoveal locations. These were resized to  $1.57^\circ \times 2.29^\circ$  so they were the same height as the other stimuli. In order to match the width of these face stimuli to the other distractor images, faces were presented on a grey rectangle which was the same width as the other stimuli in the same position (i.e.  $3.44^\circ \times 2.29^\circ$ ).

**Procedure.** The task was similar to Experiment 1a and 1b, with the following exceptions. Across all trials, participants were given the same affective search goal instruction as in Experiments 1a and 1b, but they were also given verbal instructions that some of the images may be emotional faces, predatory or poisonous animals, or dangerous objects. It was left deliberately vague what exactly these images would be, and whether these were the only threat-related images presented.

At the start of each trial participants were given a 400ms saying cue “Threat-related” to prompt the start of the trial. This was followed by the nine frame RSVP stream, with three



objects, three animals, and three faces. One of the images could be the target which could appear at position five, six, seven, or eight in the stream. On half the trials the threat-related target was present, on the other half absent. On absent trials, the target position was replaced by a neutral image; the replacement neutral stimulus was selected so that there were always three of each of the different conceptual categories (object, animal, face). All the different targets appeared equally in each experimental condition, and all other within-subject's variables were counterbalanced. At the end of the RSVP stream, a screen with a "?" appeared, with participants then indicating whether they believed a threat related image had been presented on that trial, using the 'c' and 'm' keys, with the response-answer association counterbalanced between participants. In total there were three blocks of 64 trials. At the start of the task, participants completed an eight-trial practice block which used the same stimuli as the practice block in Experiment 1a and 1b.

### **Results and Discussion**

Unlike Experiment 1a and 1b, A-prime ( $A'$ ) detection sensitivity index was the dependent variable (see Table 2), rather than accuracy, due to the use of the present/absent response rather than identification response.  $A'$  is a non-parametric analogue of  $d'$ , computed using hit rate and false alarm rate from the present/absent task response (Stanislaw & Todoroff, 1999; Zhang & Mueller 2005).  $A'$  ranges from .5, which indicates that a signal cannot be distinguished from noise, to 1, which corresponds to perfect performance. This measure removes potential response bias which can influence binary response measures such as the present/absent judgment.

A 2×2 repeated measures ANOVA was conducted using  $A'$  as the dependent variable, and distractor conceptual category (face/ scene) and distractor affect (neutral/ threat-related) as the within-participants factors. This revealed that there was no significant difference between the two distractor conceptual categories,  $F(1,15) < .01$ ,  $p = .980$ ,  $\eta_p^2 < .01$ . Further,

there was no difference between the detection sensitivity on trials when the distractor was threat-related compared to when it was neutral,  $F(1,15) = .46, p = .456, \eta^2_p = .04$ .

Additionally, against our original hypothesis, there was no interaction between the distractor conceptual category and distractor emotion,  $F(1,15) = .08, p = .785, \eta^2_p = .01$ .

**Table 2.** Mean and standard deviations for  $A'$  detection sensitivity across all distractor and search conditions within Experiment 2 and 3. See supplementary materials 1 for accuracy data for both experiments.

		Fear face distractors	Neutral face distractors	Mutilation scene distractor	Neutral scene distractors
Experiment 2 ( $n = 16$ )	Threat search	.75 (.07)	.75 (.08)	.74 (.07)	.76 (.08)
Experiment 3 ( $n = 24$ )	Fear face search	.79 (.08)	.83 (.03)	.85 (.03)	.84 (.03)
	Mutilation search	.82 (.05)	.83 (.02)	.78 (.08)	.84 (.02)

To determine whether the null finding between the threat-related and neutral distractors was sensitive, we conducted Bayesian pairwise comparisons. The prior in this case was  $A' = .10$ , which was the largest  $A'$  raw effect size taken from a similar task using fearful faces as a search goal (Brown et al., in press). This revealed that, as in Experiment 1a and 1b, there was no difference between the detection sensitivity of threat-related scenes relative to neutral scenes,  $t(15) = .59, p = .567$ , Hedges'  $g_z = .14, B_{H[0, .10]} = .34$ .<sup>4</sup> Further, the Bayes factors revealed that there was no difference between the fearful faces and the neutral faces,

<sup>4</sup> For convenience, the stopping rule was checked using Zoltan Dienes online calculator (Dienes, 2008) which produced a Bayes factor of .33 for the fearful face distractor versus the neutral face distractor in the threat-related scene search. The subsequent analyses were computed using an R code version of Baguley and Kaye's (2010) calculator, and produced a Bayes factor of .338, and was thus rounded up to .34, hence the difference with this Bayes factor and the stopping rule of .33 and below.

$t(15) = .73, p = .479$ , Hedges'  $g_z = .18$ ,  $B_{H[0, .10]} = .17$ . These null results are particularly striking given that the fearful faces were part of the target search category.

The intentions of Experiments 1 and 2 were to provide a test of whether a broad search goal for an affective category could induce involuntary attentional capture by congruent threat-related distractors. Despite assessing the evidence for goal-induced attentional capture across multiple threat-related categories, distractor dimensions, response types, and even when these images were congruent with one of the subordinate target categories of threat, the results were consistent in failing to detect any generalisation of goal-driven capture by threat-related distractors.

### **Experiment 3**

The lack of evidence for any goal-driven attentional capture by threat-related stimuli, even when they were part of the attentional set, strikingly contrasts with our previous findings using a similar paradigm (Brown et al., in press; 2018a; 2018b). Within these previous investigations, we found that searching for a single conceptual category of affective stimuli (i.e. a category based on shared properties above and beyond the affective properties, e.g. 'threatening animals') consistently resulted in strong attentional capture by other affective stimuli from the same category, even when they appeared in task-irrelevant locations. This could suggest that goal-driven attentional capture cannot generalise across semantically heterogeneous categories based on shared affective properties alone. However, it was important to confirm that the distractor stimuli used in the present study are capable of eliciting robust goal-driven capture effects when the search goals are restricted to a single sub-category of affective stimuli. To this end, we conducted a final experiment in which participants were given search goals for each of the two distractor subcategories (fearful faces and the mutilation scenes) in separate blocks. We predicted that these threat-related

categories would only capture attention, relative to the neutral stimuli, when participants were searching for that specific category of threat in the central RSVP stream.

## Methods

**Participants.** Participants were recruited until all Bayes factors for the pairwise comparisons between neutral and threat-related distractors were sensitive, across all distractor categories in both search goal conditions. The Bayes factors were checked at the end of each day of testing. This led to 24 participants being recruited, of which 18 were female and 6 were male, the average age of which was 23.50 ( $SD = 3.49$ ). Participants were remunerated with course credits or a small cash payment. A power analysis using G\*power software revealed that our final sample had the power of  $\beta = .80$  to detect a pairwise effect of  $d_z = .60$  or above ( $\alpha = .05$ ; Faul et al., 2007).

**Stimuli and procedure.** The task was identical to Experiment 2 with the exception of the following changes. In order to experimentally manipulate the specificity of search goal, participants were instructed to search for fearful faces or scenes of mutilation and death in different blocks. In total there were six blocks which were made up of 64 trials, with three blocks where participants searched for “Fearful faces” and three where they searched for “Injury and death”. These blocks were presented in a mixed order (i.e. Fear – Mutilation – Fear – Mutilation – Fear - Mutilation) with the order counterbalanced between participants. Before each block, participants were instructed what the upcoming target category would be, requiring them to press the space bar to continue to the next block. Additionally, before each trial began participants would be prompted with a 400ms text warning for the category, either “Fearful faces” or “Injury and death”. The face target set consisted of 12 faces, meaning that in addition to the eight faces used in Experiment 2, four additional fearful face targets were added from the NimStim and Amsterdam Dynamic Facial Expression Set (ADFES; Van Der

Schalk, Hawk, Fischer & Doosje, 2011). The stimulus set for scenes of mutilation and death consisted of 12 images that were taken from the IAPS and the GAPED stimulus sets.

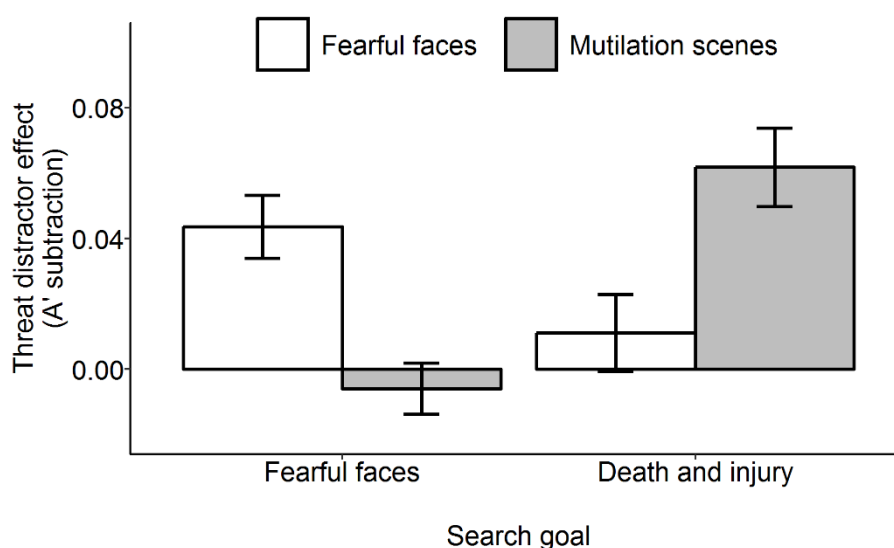
Additionally, due to the inclusion of mutilation scenes in the central stream we replaced the neutral animal filler images with neutral human scene filler images, thus matching the experiments for the semantic relations between the filler and the target categories. These neutral scene images consisted of 140 images, sourced from the IAPS and NAPS image sets, or from Google images. Eighty images sourced from Google were selected based on their similarity to the images from the IAPS and NAPS; they included scenes of people shopping, on public transport, or at work. The neutral objects consisted of 128 neutral objects, which were the same as in Experiment 2, though 8 were removed. The neutral faces consisted of 140 faces, with four male and four female faces from the lifespan database added to those used in Experiment 2 (Minear & Park, 2004). The stimuli presented in distractor locations were identical to those presented in the previous experiment.

### **Results and Discussion**

We conducted a  $2 \times 2 \times 2$  repeated measures ANOVA on the  $A'$  score (see Table 2), using search goal (fearful faces/ death and injury), distractor conceptual category (faces/ scenes), and distractor valence (neutral/ threat-related) as factors. This revealed that there was a significant effect of search goal,  $F(1,23) = 5.66$ ,  $p = .026$ ,  $\eta^2_p = .20$ , with participants more accurately detecting the fearful faces than the scenes of death and injury. There was also a marginally significant effect of distractor conceptual category,  $F(1,23) = 3.32$ ,  $p = .081$ ,  $\eta^2_p = .13$ , whereby face distractors resulted in lower detection sensitivity relative to the scene distractors.

The category of distractor did not significantly interact with the valence of the distractor,  $F(1,23) = .01$ ,  $p = .956$ ,  $\eta^2_p < .01$ . However, the current search goal of the participant did significantly interact with the category of distractor presented,  $F(1,23) =$

10.48,  $p = .001$ ,  $\eta^2_p = .37$ , with both scenes and face distractors resulting in lower detection sensitivity when congruent with the current search goal, relative to when they were incongruent. Current search goal also marginally interacted with the valence of the distractors,  $F(1,23) = 4.10$ ,  $p = .055$ ,  $\eta^2_p = .15$ , such that when participants were searching for scenes of injury and death, participants were worse at detecting the target when the distractor was threat related relative to neutral. Importantly, both of these interactions were driven by a significant three-way interaction between current search goal, distractor conceptual category, and distractor threat-relevance,  $F(1,23) = 15.71$ ,  $p = .001$ ,  $\eta^2_p = .41$ . As can be seen in Figure 2, this interaction reflected interference from the threat-related (versus neutral) distractors only when these matched the current threat-related category being searched for.



**Figure 2.** Mean threat-related distractor effect as a function of distractor category and search goal. The threat-related distractor effect reflects the decrease in target detection sensitivity ( $A'$ ) in the presence of a threat-related distractor versus a matched neutral distractor. Larger distractor effects depict a greater decrement in target detection sensitivity. Error bars reflect within-subjects standard error.

In order to break down the three-way interaction, we conducted four Bayesian pairwise comparisons between  $A'$  when the distractor was threat-related compared to its matched neutral counterpart, within each search goal condition. This revealed that when participants were searching for the threat-related scenes, the detection sensitivity did not

differ between fearful face distractors and the neutral face distractor,  $t(23) = 1.25$ ,  $p = .224$ , Hedges'  $g_z = .25$ ,  $B_{H[0, .10]} = .33$ . Additionally, the threat-related scene did not differ from neutral scenes in their influence on detection sensitivity of the fearful faces,  $t(23) = .88$ ,  $p = .386$ , Hedges'  $g_z = -.18$ ,  $B_{H[0, .10]} = .04$ . Therefore, when incongruent with the current search goal, there was no evidence of attentional capture by threat-related distractors. When, however, the distractors were congruent with the current search goal, the threat-related distractors produced a significant decrement in detection sensitivity relative to their neutral counterpart – reflecting goal-driven attentional capture. This was true for both faces and the mutilation scenes,  $t(23) = 3.36$ ,  $p = .003$ , Hedges'  $g_z = .68$ ,  $B_{H[0, .10]} = 46.13$ ;  $t(23) = 3.85$ ,  $p = .001$ , Hedges'  $g_z = .78$ ,  $B_{H[0, .10]} = 265.91$ , respectively.

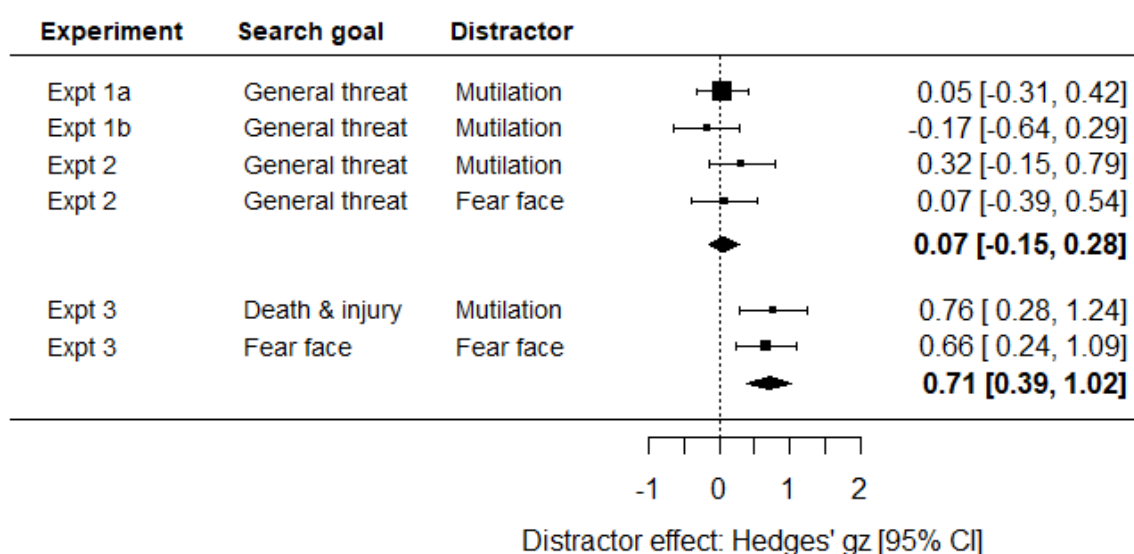
### Internal Meta-analysis

As well as analysing each experiment individually, we also chose to meta-analytically compute the overall threat distractor effect size across experiments depending on the breadth of the search goal. We initially conducted a moderation analysis comparing the cumulative Hedges'  $g_z$  standardised effects of the threat distractor (i.e. neutral minus threat distractor accuracy) when participants were searching for a broad category of threat consisting of multiple subcategories ( $k = 4$ ), compared to conditions where participants were searching for a specific category of threat and the distractors were also from the same specific category ( $k = 2$ ). The full details of this analysis are reported in supplementary materials 2.

The moderation analysis revealed that across experiments there was a significantly larger distractor effect when participants held a search goal for a specific category of threat compared to a broad affective category, Hedges'  $g_z = .64$ ,  $p = .001$ , 95% CI[.26, 1.02]. To follow-up this difference, we computed two fixed effects meta-analyses within each search goal condition (broad/specific; see Figure 3). These revealed that there was a near zero and non-significant decrement in accuracy incurred by threat distractors in the broad search

condition, Hedges'  $g_z = .07$ ,  $p = .599$ , 95% CI[-.15, .28]. The Bayes factors revealed strong evidence favouring the null hypothesis with both the effect size predicted at the start of the investigation,  $B_{H[0, .8]} = .13$ , as well as a reduced prior effect size which should be sensitive to even small effects,  $B_{H[0, .4]} = .26$ .

When, however, we focused on the two conditions in Experiment 3, where the distractor matched the single subcategory of threat being searched for, there was a large significant effect, Hedges'  $g_z = .71$ ,  $p < .001$ , 95% CI[.39, 1.02], and strong evidence favouring the experimental hypothesis,  $B_{H[0, .8]} = 425249.40$ . The cumulative evidence therefore matched the evidence at the individual experiment level: There was evidence against goal-driven attentional capture when participants held a broad search goal for multiple threat-related stimuli, but strong evidence of goal-driven attentional capture within a single affective-conceptual category.



**Figure 3.** Forest plot presenting the Hedges'  $g_z$  standardised effect sizes and 95% confidence intervals from across all conditions where the distractor was threat-related. Fixed effects cumulative effect sizes are presented as diamonds. These were computed based on whether the current search goal for threat was broad (top diamond) or specific (bottom diamond).



### General Discussion

Across our first three experiments, we consistently found that asking participants to adopt a broad attentional set for threat did not induce attentional capture by threat-related distractors (fearful faces and threat-related scenes), despite the clear affective overlap between goal and distractor. No goal-driven capture was observed even when a third of the targets which made up this search goal were drawn from the same sub-category as the distractors; that is, when participants searched for fearful faces, threatening objects, and threatening animals, fearful faces did not capture attention. However, these same affective distractors produced clear attentional capture effects when participants searched for the same single sub-category of threat-related stimuli.

The absence of goal-driven capture across the general category of ‘threat’, defined purely by common affective properties among category items, appears to contrast with previous demonstrations of robust goal-driven capture by conceptual categories (Wyble et al., 2013; Brown et al., 2018a; 2018b; in press; Wu, Pruitt, Zinszer & Cheung, 2017; Nako, Wu, Smith & Eimer, 2014). However, in previous examples of goal-driven capture by conceptual categories, the category was defined by semantic and functional relations between items rather than just affective associations (e.g. dangerous objects also have the common semantic property of being ‘objects’). Indeed, in Experiment 3 when participants searched for single conceptual categories of threat, which held common semantic as well as affective properties (e.g. ‘fearful faces’), we found reliable goal-driven capture. This implies a distinction in the relation of goal-driven attention to categories defined by non-affective (e.g. perceptual or semantic) versus purely affective properties – while the former can induce goal-driven capture, the latter cannot. One might speculate that this reflects differences in the ability of affective information, versus perceptual or semantic information, to influence goal-driven attention.

An alternative possibility, however, could lie in suggestions that involuntary capture can only result from prioritisation of a single perceptual feature (e.g. colour or shape; van Mooreselaar, Theeuwes & Olivers, 2014), or only for multiple features related to a single object file (e.g. Berggren & Eimer, 2018). It might be argued that affective categories are, therefore, too perceptually heterogeneous to induce goal-driven attentional capture. By this account, previous demonstrations of goal-driven capture by conceptual categories are argued to reflect individuals tuning attention towards a limited set of features which were typical of a category (Evans & Treisman, 2005). In support of a key role of perceptual features in category search, Yu, Maxfield and Zelinsky (2016) determined the prevalence of category-consistent features across stimuli using computational techniques, and found that participants were faster to detect a target from a category if it contained more of these consistent features. Further, participants are faster to detect targets from categories with a lot of similar features compared to those that are more varied (Hout, Robbins, Godwin, Fitzsimmons & Scarince, 2017).

Such an account would also explain why, when participants were searching for a single category of threat in Experiment 3, another category of threat did not capture attention. This initially appears to contradict previous research which has revealed that when individuals are searching for a single exemplar from a category, other exemplars from the same category capture attention despite not being the specific target. For instance, research has found that searching for a single item of clothing (e.g. jeans) resulted in capture by other clothes (e.g. t-shirts; Nako et al., 2014) and searching for one type of animal (e.g. a fish) led to capture by another animal (e.g. a bird; Telling, Kumar, Meyer & Humphreys, 2010). When, however, participants were searching for either fearful faces or mutilation, the other category did not capture attention, despite being threat related. If, however, previous generalisation effects were due to attentional capture by common features across exemplars

within a category, then the previous instances of conceptual attentional capture could simply be due to visual overlap between stimuli. For instance, in Telling and colleagues' experiment, birds and fish were the only stimuli to have facial features, and in Nako and colleagues' work the t-shirts and jeans were the only stimuli to have stitching and creases. Fearful faces and mutilation images may be visually distinct enough that the participants could search for one without inducing capture by the other threat category.

On the other hand, the suggestion that goal-driven capture can be induced only by perceptual features is challenged by Giammarco and colleagues' (2016) finding of goal-driven capture by a heterogeneous artificial category. In this case, participants were trained on an artificial category learning task until they could report each object's category at 90% accuracy. The participants then searched for these heterogeneous categories in an RSVP task. It was found that a task-irrelevant distractor only interfered with target detection if it was from the same artificial category as the current search goal, even though these categories had only been learnt in the same session. It might be argued, however, that the extensive training enabled participants to learn common features within the category - thus a specific feature-based account cannot be fully discounted. Further, there is the possibility that conceptual categories with more complex relationships beyond simple group membership are searched for differently from artificially grouped objects (Wu & Zhao, 2017). Future research is required to delineate the role of experience and the depth of relationship between the category exemplars in categorical attentional search tasks.

Regardless of the specific mechanism involved, the present work highlights an important boundary condition for goal-driven attention that informs theoretical understanding of attentional capture by affective stimuli. A growing body of research has pointed to a goal-driven account of attentional capture by affective stimuli, highlighting that this capture often depends on task-relevance rather than occurring unconditionally (e.g. Vromen et al., 2016;

Lichtenstein-Vidne et al., 2012; Stein et al., 2009). Our current results clearly demonstrate that relevance to an affective top-down attentional goal is not wholly sufficient to induce goal-driven attentional capture; if this were the case then participants should have been distracted by all threat-related stimuli when adopting a search goal for this category. Instead, the results suggest that goal-driven capture by irrelevant threat-related stimuli may only occur when that specific conceptual category of threat-related stimuli (e.g. fearful faces or mutilation scenes) is adopted as an attentional goal.

Indeed, we propose that in order for involuntary attentional capture to emerge there has to be the intention to search for that specific stimulus category, as well as knowledge of which specific features define that category, not just the relevance of the affective distractor to the current task aims. In light of the present findings, it appears possible that previous demonstrations of relevance effects could be explained by relevance cueing participants to adopt a top-down search goal for the specific conceptual category of threat-related stimuli. For example, in Lichtenstein-Vidne and colleagues' (2012) study, participants may have noticed that the targets were often threat-related scenes and adopted a more specific attentional goal for these scenes because of their motivational relevance, even though this strategy was not necessary to complete the task.

We note that while we did not observe involuntary goal-driven effects across affective categories, participants nevertheless appeared capable of using these categories to direct voluntary goal-driven attention (i.e. they could detect or identify different stimuli from across this visually heterogeneous affective category). An item level analysis revealed that each participant correctly reported the majority of the threat-related target exemplars at least once across the experiment (Expt 1a  $M = 93\%$ ; Expt 1b  $M = 90\%$ ; Expt 2  $M = 94\%$ ). The data therefore suggest that participants perceived stimuli drawn from multiple semantic categories (e.g. faces, animals, objects) as matching the search goal, meaning that, at some level,

participants were able to hold multiple/broad threat-related categories active as a search goal without inducing involuntary attentional capture.

In this respect, our results are intriguing in suggesting that the criteria for a stimulus category to guide voluntary attention are not the same as those which produces involuntary goal-driven effects. Our results could perhaps be accommodated within models of hybrid visual search (e.g. Cunningham & Wolfe, 2014; Wolfe, 2012), which propose that when individuals only have to search for a limited set of visual features to complete a task, these features are prioritised as a search template. When, however, individuals are required to search for multiple visually distinct targets they adopt a different strategy relying on analysing objects at a post-selective level. Hybrid visual search theory has not yet been explored in the context of involuntary goal-driven capture; however, one might speculate that involuntary effects occur only in contexts where templates for clearly defined stimulus features can be adopted, and not in contexts where the later strategy is called upon.

The current experiments were designed to investigate a within-participants experimental question, rather than individual differences in attentional capture. However, the specificity of goal-driven attentional capture seen in the present investigation results appears to parallel patterns seen in recent anxiety disorder research. A recent meta-analysis revealed that across a range of anxiety disorders (e.g. phobias, Post-Traumatic Stress Disorder, social anxiety), specific sub-categories of threat-related stimuli which were particularly relevant to that disorder (e.g. angry faces in social anxiety) captured attention more than other threat-related stimuli (Pergamin-Hight, Naim, Bakermans-Kranenburg, van Ijendoorn & Bar-Haim, 2015). These rather specific clinical attentional biases, along with the present evidence, are consistent with the notion that individuals assume search goals only for the specific threat-related stimuli which are relevant to their current concerns.

An interesting question is how such a specific goal-driven form of attentional capture could work in a real-world setting, where participants would be unlikely to search for a specific threat across all situations at the cost of all other search goals. Here, we propose that participants may use prior knowledge and learned associations between a context and the objects that may appear there, in order to determine what specific category to search for. A range of different threat-related stimuli might be deemed important at one time, but contextual cues would allow prioritisation of specific categories of threat which are more likely to appear in a specific moment. For example, a dark inner-city alleyway would be more likely to cue a goal for potential muggers than poisonous snakes, while a hike in Joshua Tree National Park might do the opposite.

Within the current investigation, we explored how a search goal for threat could induce attentional capture, without the need to induce an internal state or select for personality differences that have previously been required to elicit attentional capture by threat (e.g. Bar-Haim et al., 2007). It is highly likely, however, that many real-world contexts in which an individual would be expecting and searching for threat would also be highly arousing. To fully understand the real-world operation of goal-driven attention, future work should therefore explore potential interactions between current arousal or anxiety with affective search goals.

In conclusion, we found no evidence that involuntary attentional capture by threat-related stimuli can be induced by a broad affective search goal, despite the same stimuli eliciting robust goal-driven capture when a more specific affective search goal was adopted. This finding highlights an important boundary condition for goal-driven attentional capture, which must be considered in any goal-driven account of attentional capture by affective stimuli.

### Open Practices Statement

All tasks, unlicensed images sourced or created for this investigation, as well as data and analysis scripts are available via the Open Science Framework: [osf.io/ju87s](https://osf.io/ju87s).

### References

- Bacon, W.F., & Egeth, H.E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*, 485-496.
- Baguley, T., & Kaye, W.S. (2010). Review of “Understanding Psychology as a Science: An Introduction to Scientific and Statistical Inference”. *British Journal of Mathematical & Statistical Psychology*, *63*, 695–698.
- 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*, 1-24.
- Berggren, N., & Eimer, M. (2018). Object-based target templates guide attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 1368-1382.
- Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences*, *11*, 307-316.
- Brown, C. R. H., Berggren, N. & Forster, S. (in press). Testing a goal-driven account of attentional capture by threat. *Emotion*
- Brown, C. R. H., Duka, T., & Forster, S. (2018a). Attentional capture by alcohol-related stimuli may be activated involuntarily by top-down search goals. *Psychopharmacology*, *235*, 2087–2099.
- Brown, C. R. H., Forster, S., & Duka, T. (2018b). Goal-driven attentional capture by appetitive and aversive smoking-related cues in nicotine-dependent smokers. *Drug and Alcohol Dependence*, *190*, 209-215.
- Carretie, L. (2014). Exogenous (automatic) attention to emotional stimuli: A review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*, 1228-1258.
- Cisler, J.M., Bacon, A.K., & Williams, N.L. (2009). Phenomenological characteristics of attentional biases towards threat: A critical review. *Cognitive Therapy and Research*, *33*, 221-234.

- Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*(4), 1585-1599.
- Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, *43*(2), 468-477.
- Dienes, Z. (2008). Understanding psychology as a science: An introduction to scientific and statistical inference. Hampshire, England: Palgrave Macmillan. Retrieved from [http://www.lifesci.sussex.ac.uk/home/Zoltan\\_Dienes/inference/](http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/inference/)
- Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, *6*, 274-290.
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(781), 1-17.
- Dienes, Z. (2016). How Bayes Factors change scientific practice. *Journal of Mathematical Psychology*, *72*, 72-89.
- Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un) consciousness. *Annual Review of Psychology*, *61*, 467-490.
- Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free?. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476-1492.
- Everaert, T., Spruyt, A., & De Houwer, J. (2013). On the malleability of automatic attentional biases: Effects of feature-specific attention allocation. *Cognition & Emotion*, *27*, 385-400.
- 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.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030-1044.



- Folk, C. L., Leber, A. B., & Egeth, H. E. (2002). Made you blink! Contingent attentional capture produces a spatial blink. *Attention, Perception & Psychophysics*, *64*, 741-753.
- Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 317-329.
- Giammarco, M., Paoletti, A., Guild, E. B., & Al-Aidroos, N. (2016). Attentional capture by items that match episodic long-term memory representations. *Visual Cognition*, *24*(1), 78-101.
- Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set for displaywide visual features. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 699-706.
- Grimshaw, G. M., Kranz, L. S., Carmel, D., Moody, R. E., & Devue, C. (2018). Contrasting reactive and proactive control of emotional distraction. *Emotion*, *18*(1), 26-38.
- Hout, M. C., Robbins, A., Godwin, H. J., Fitzsimmons, G., & Scarince, C. (2017). Categorical templates are more useful when features are consistent: Evidence from eye movements during search for societally important vehicles. *Attention, Perception, & Psychophysics*, *79*(6), 1578-1592.
- Jeffreys, H. (1961). Theory of probability (Eds.), *Oxford Classic Texts in the Physical Sciences*. Oxford University Press: Oxford.
- Kennedy, B.L., & Most, S.B. (2015). The rapid perceptual impact of emotional distractors. *PloS One*, *10*, e0129320.
- Klinger, E. (2013). Goal commitments and the content of thoughts and dreams: Basic principles. *Frontiers in Psychology*, *4*(415), 1-17.
- Koster, E. H., 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 & Therapy*, *42*(10), 1183-1192.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings*. NIMH Center for the Study of Emotion and Attention, 39-58.

- Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*(1), 132-138.
- Lichtenstein-Vidne, L., Henik, A., & Safadi, Z. (2012). Task relevance modulates processing of distracting emotional stimuli. *Cognition & Emotion*, *26*, 42-52.
- Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596-610.
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*, 630-633.
- Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review*, *12*, 654-661.
- Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans? *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1283-1288.
- Öhman, A. (1992). *Attention and information processing in infants and adults: Perspectives from human and animal research*. Hillsdale, NJ: Erlbaum.
- Öhman, A. (2005). The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology*, *30*, 953-958.
- Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: a meta-analysis. *Clinical Psychology Review*, *35*, 10-18.
- Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Reviews in the Neurosciences*, *14*(4), 303-316.
- Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition & Emotion*, *29*(4), 687-694.

- Shang, H., & Bishop, I. D. (2000). Visual thresholds for detection, recognition and visual impact in landscape settings. *Journal of Environmental Psychology*, *20*(2), 125-140.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.
- Stein, T., Zwickel, J., Ritter, J., Kitzmantel, M., & Schneider, W. X. (2009). The effect of fearful faces on the attentional blink is task dependent. *Psychonomic Bulletin & Review*, *16*, 104-109.
- Telling, A. L., Kumar, S., Meyer, A. S., & Humphreys, G. W. (2010). Electrophysiological evidence of semantic interference in visual search. *Journal of Cognitive Neuroscience*, *22*(10), 2212-2225.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*, 77-99.
- Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, *1*(1): 29, pp. 1-15,
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242-249.
- Van Der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, *11*(4), 907-920.
- van Moorselaar, D., Theeuwes, J., & Olivers, C. N. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention?. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1450-1465.
- Vervoort, E., Vervliet, B., Bennett, M., & Baeyens, F. (2014). Generalization of human fear acquisition and extinction within a novel arbitrary stimulus category. *PloS one*, *9*(5), e96569.
- Vogt, J., De Houwer, J., Koster, E. H. W., Van Damme, S., & Crombez, G. (2008). Allocation of spatial attention to emotional stimuli depends upon arousal and not valence. *Emotion*, *8*(6), 880-885.

- Vogt, J., De Houwer, J., Crombez, G., & Van Damme, S. (2013). Competing for attentional priority: Temporary goals versus threats. *Emotion, 13*(3), 587-598.
- Vromen, J. M., Lipp, O. V., Remington, R. W., & Becker, S. I. (2016). Threat captures attention, but not automatically: Top-down goals modulate attentional orienting to threat distractors. *Attention, Perception, & Psychophysics, 78*, 2266-2279
- Wells, A., & Matthews, G. (1994). *Attention and emotion: A clinical perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wu, R., & Zhao, J. (2017). Prior knowledge of object associations shapes attentional templates and information acquisition. *Frontiers in Psychology, 8*(843), 1-6.
- Wu, R., Pruitt, Z., Zinszer, B. D., & Cheung, O. S. (2017). Increased experience amplifies the activation of task-irrelevant category representations. *Attention, Perception, & Psychophysics, 79*(2), 522-532.
- Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 861-871.
- Yu, C. P., Maxfield, J. T., & Zelinsky, G. J. (2016). Searching for category-consistent features: A computational approach to understanding visual category representation. *Psychological Science, 27*(6), 870-884.
- Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70*(1), 203-212.