CAN EMOTIONAL SINGLETONS BE SUPPRESSED?

by

Minwoo Kim

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychological and Brain Sciences.

Summer 2021

CAN EMOTIONAL SINGLETONS BE SUPPRESSED?

by

Minwoo Kim

Approved: ____
Tania Roth, Ph.D.
Chair of the Department of Psychological & Brain Sciences

Approved: ____
John Pelesko, Ph.D.
Dean of the College of Arts & Sciences

Approved: ____
Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

James E. Hoffman, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Charles L. Folk, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Jared Medina, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Philip A. Gable, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Timothy Vickery, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I would first like to thank my advisor, Dr. James E. Hoffman, who was committed to providing good mentorship during the entire process. Your insightful feedback brought the dissertation to a higher level and sharpened my reasoning.

I would like to thank my colleague Hansol Rheem for sharing his expert knowledge on conducting and analyzing Mouse-tracking experiments, which has been extremely useful in writing chapter 4.

In addition, I would like to thank my cognitive cohort Alyssa Lompado and Irem Yildirim for moral support and my parents for believing in my dreams.

Finally, I could not have completed this dissertation without the support of my beloved wife Younghwa Kim who is my best friend, academic colleague, but most importantly my soulmate.

TABLE OF CONTENTS

Chapter

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Emotional stimuli can capture our visual attention even when we are trying to ignore them. However, recent research shows that even highly physically salient stimuli can be prevented from capturing attention when presented at a suppressed location. The primary purpose of the current study was to test whether emotional pictures automatically capture attention and are capable of breaking through attentional suppression.

Attention capture and suppression was tested with a paradigm using statistical learning with the additional singleton procedure (Wang & Theeuwes, 2018a, b). Participants discriminated a target in a display containing a physically salient, singleton distractor image that was either emotional or neutral. Distractor pictures were much more likely to occur in one location compared to other locations which proactively suppressed objects appearing in that location.

RT results show that both emotional and neutral singletons appearing at non-suppressed, low probability locations captured attention with larger effects for emotional than neutral distractors. Both distractors appearing at the high probability location were suppressed, showing smaller interference effects, but emotional distractors continued to produce more interference than neutral. However, mouse tracking data indicates that the emotional interference effect in the high location was actually not due to spatial capture of attention. Spatial capture in the high location was actually suppressed but more time was consumed to suppress an emotional distractor compared to a neutral one. This suggests spatial attention capture by emotional salience is not automatic and can be completely suppressed.

Keywords: emotional salience, attention capture, additional singleton, suppression, automaticity, mouse-tracking

Chapter 1

GENERAL INTRODUCTION

Attention is usually under out volitional control, allowing us to pay attention to any object whenever we please. But a salient external event, such as a flash of light or a sudden movement, can draw our attention against our will. The involuntary recruitment of attention caused by a physically salient property of an external stimulus is called *attention capture*. Currently, there is a consensus that a few types of physically salient visual features such as color, motion, orientation, and size can be highly effective at capturing attention (see Wolfe, & Horowitz, 2017 for review). In fact, capture by physically salient visual objects has been considered automatic, because it occurs very *rapidly,* and *without attention* (Yantis, & Jonides, 1984). However, it has recently been shown that attention capture can be prevented by actively suppressing the processing of the salient stimulus (Sawaki, & Luck, 2010; Wang & Theeuwes, 2018a, b).

Similar to physically salient stimuli, emotionally salient stimuli can capture attention as well. This is the case for a wide variety of emotional stimuli, including emotional facial expressions (Bannerman, Temminck, & Sahraie, 2012; Grose-Fifer, Rodrigues, Hoover, & Zottoli, 2013; Padmala, Sambuco, Codispoti, & Pessoa 2018; Vuilleumier, & Schwartz, 2001), neutral stimuli conditioned with punishment and/or reward (Anderson, Laurent, & Yantis, 2011a; Schmidt, Belopolsky, & Theeuwes, 2015; Smith, Most, Newsome, & Zald, 2006; Watson, Pearson, Wiers, & Le Pelley, 2019), pictures of threatening animals (Vromen, Lipp, Remington, & Becker, 2016;

Yamaguchi, & Harwood, 2015), erotic images (Most, Smith, Cooter, Levy, & Zald, 2007), and emotional words (Arnell, Killman, & Fijavz, 2007; Hinojosa et al., 2015; Huang, Baddeley, & Young, 2008). However, whether emotional stimuli capture attention *automatically* is still an ongoing issue. While some argue that humans are 'hard-wired' to automatically attend to emotional stimuli (e.g., Ledoux, 1998; Öhman, 2009; Öhman, Flykt, & Esteves, 2001), others argue that emotional capture is not automatic (Glickman, & Lamy, 2018; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002). Also, currently there is a lack of evidence on whether emotional distractors are subject to suppression. The purpose of the current study is to provide additional evidence on whether emotional attention capture can be suppressed. The following sections discuss background theories and related evidence regarding the mechanism of attention capture in both general and emotional contexts.

1.1 Original Theories of Attention Capture

To understand how emotional distractors capture attention, it is necessary to understand the general theory of attention capture. In the classical model of visual attention, low level sensory information of visual stimuli such as color or orientation are extracted by an early pre-attentive processing stage that operates without limitations in attentional capacity. In order for the visual information to reach post-attentive working memory and awareness stages, limited capacity attentional selection mechanism must reduce the incoming information to one or a small number of objects(Broadbent, 1958; Treisman, 1964). Theories that focused on the mechanism of selection heavily debated whether a physically salient singleton (a stimulus containing a unique feature in an otherwise homogenous set of objects) captures attention via an automatic bottom-up mechanism or through top-down control (see Theeuwes, 2010 for review). The *bottom-up saliency* account (also called stimulus-driven capture; Theeuwes, 1992) argues that physically salient distractors always capture attention automatically and ultimately reach working memory. Volitional control over attention is subsequently regained only after attention is disengaged from the singleton distractor.

Stimulus-driven capture has been typically observed with the *additional singleton paradigm*, in which participants visually search for a unique target among a homogenous display of items (e.g., a green diamond shaped target among green circles). When one of the task-irrelevant distractor items (e.g., the circles) is a singleton distractor with a physically salient task-irrelevant feature (e.g., a red circle among green circles), the singleton distractor automatically captures attention and delays responses to the target.

The opposing goal-driven account, called *contingent capture* (Folk, Remington, & Johnston, 1992), suggests that capture occurs because the salient feature of the singleton distractor is relevant to the participant's current top-down goals. For example, Folk and colleagues demonstrated that color singleton distractors only captured attention when the target was also defined by color. However, abrupt onset singleton distractors did not capture attention when the target was defined by color. The reverse was also true: onset singletons captured attention when targets were also defined by onset, but color singletons did not. Bacon and Egeth (1994) also demonstrated how top-down mechanisms could account for attention capture by salient singletons that do not share features with the target. They suggested that participants can adopt an implicit top-down strategy called *singleton detection mode,* when the target and the distractors are both singletons but in different dimensions. For example, if people search for a shape singleton target, such as a green triangle among green circles, presentation of a color singleton distractor, such as a red circle, would capture attention because participants are tuned in to find a singleton that stands out from the rest of the display. In contrast, if the target is no longer a singleton and is presented in a heterogenous display (e.g., red circle among green, yellow, blue, and purple circles) participants may adopt a *feature-search* mode (e.g., search for something red instead of something that stands out) in which the salient singleton would no-longer capture attention.

1.2 The Integrated Priority Map Framework

One of the main problems with the original debate was that neither of the two accounts succeeded in fully explaining all the capture phenomena. One example is *priming of pop-out* (Maljkovic, & Nakayama, 1994), in which processing of a stimulus feature is enhanced when the stimulus is presented as a target feature on the previous trial. Priming of pop-out is a frequently observed confounding variable when studying goal-driven attention. For example, when a target is a green circle and the distractor is a red circle in the current trial, the red circle will capture attention if it was a target in the previous trial. In other words, the history of previous selections can affect attention capture, even when it contradicts the current goals.

In the last decade, the *integrated priority map theory* (Awh, Belopolsky, & Theeuwes, 2012; Theeuwes, 2019; also see Luck, Gaspelin, Folk, Remington, & Theeuwes, 2021 for the latest update) was proposed in order to provide a working theoretical framework that incorporates most aspects of capture. This theory proposes that capture by a distractor is determined by the sum of signals generated from three mechanisms: stimulus-driven, goal-driven, and history-driven capture (see Figure 1).

History-driven capture is caused by lingering biases in our memory that are neither driven by bottom-up salience nor top-down goals (see Theeuwes, 2010). Priming of pop-out is a good example of history-driven capture since capture is driven by the history of selection that occurred in previous trials. The summed signals are then projected on to a spatial salience priority map. When multiple items are displayed, the item that received the strongest signal in the map receives attention in a winner-take-all fashion and is then admitted into working memory. Unlike the original debate where capture is strictly bottom-up or top-down, in the current framework capture is a result of contributions from all three mechanisms.

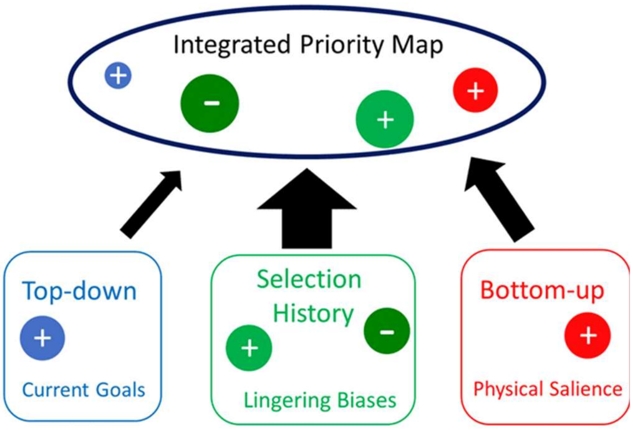


Figure 1. Illustration of the integrated priority map theory. Signals from the bottom-up (stimulus-driven), top-down (goal-driven), and selection history (history-driven) mechanisms are integrated in a spatial priority map. Display items that receive the highest summed weight captures attention in a winner-take-all way. '+' signs represent upweighting and '–' signals represent suppression of the generated signals. Adopted from Theeuwes (2018).

1.3 Active Suppression

Active suppression is another feature of selection history that has been assimilated into the integrated priority map framework which holds that suppression can occur in two major ways: reactive suppression and proactive suppression (see Geng, 2014). *Reactive suppression* occurs when a distractor first captures attention but is suppressed once it is determined not to be the target (also see *search and destroy* hypothesis: Moher, & Egeth, 2012). This means that the distractor initially captures attention and is then suppressed via a feedback loop from higher cognitive systems. On the other hand, *proactive suppression* operates at an earlier stage than reactive suppression. According to the *signal-suppression* hypothesis (Sawaki, & Luck, 2010) physically salient distractors always automatically generate a bottom-up, salience-driven 'attend-to-me' signal. The signal would naturally lead to attention capture without any intervention. However, it is possible to prevent the attend-to-me signal from affecting the priority map at an early stage.

Sawaki and Luck (2010) demonstrated proactive suppression using event-related potentials (ERP), which are useful for identifying the different stages of mental processing due to their high temporal resolution. They measured an ERP component related to attention capture called the N2pc (N2, posterior contralateral; Luck, & Hillyard, 1994) which peaks 150 to 250 ms after onset of the attention capturing distractor. They also measured the Pd component (distractor positivity; Hickey, Di Lollo, & McDonald, 2009; Sawaki, & Luck, 2010) which shares similar topographical and temporal characteristics with the N2pc but has a positive polarity and reflects suppression of distractors. Participants viewed a series of letters which were presented in two regions that were either attended or ignored across blocks. The task was to identify a target letter that was defined by its unique feature and identity (e.g., a large

“A”). One of two types of distractors were presented among the letters to measure different modes of attention capture: a target-similar distractor (e.g., a small “A”) that measured contingent top-down capture, and a physically salient color singleton distractor (e.g., red among green) to measure bottom-up capture. Results show that a contingent distractor, captured attention due to overlapping features with the target only when it was attended. This is consistent with findings from Folk and Remington (1992) that show distractors which share features with the target (i.e., distractors that matching the current attentional set) captures attention.

Critically, the evidence of suppression is shown by the ERP results from the salient singletons. According to the top-down account, the *salient* singleton distractor should *not* capture attention because the task does not involve searching for a singleton target (Bacon & Egeth, 1994). In other words, the attentional set is not tuned to singleton detection mode but rather tune to feature detection mode because the tasks involve a specific target identity feature among a heterogenous string of letters. The lack of capture should be shown as an absence of the N2pc component because the N2pc represents attention capture. But according to the stimulus-driven account suggests, the physically salient singleton distractor should always capture attention automatically, as long as it is receiving spatial attention (Theeuwes, 1992). This means that a salient singleton should produce an N2pc when attended but not when it is not attended. The ERP results do not show patterns consistent to neither the top-down nor bottom-up accounts. Instead of the N2pc, the Pd was present both when the salient singleton distractor was attended and unattended. This suggests that the salient singleton in feature-detection is actively suppressed rather than simply not capturing attention.

Recently, Wang and Theeuwes (2018a, b) showed that suppression can be directed to a specific location on the display through statistical learning. They developed an additional singleton procedure manipulates the probability of distractor appearance at a given location (see Figure 2). They presented a homogenous array of shapes (e.g., diamonds) which contained a shape singleton (e.g., a circle) which was defined as the target. The task was to report the orientation of the bar (horizontal or vertical) within the shape singleton. One of the homogenously shaped items had a physically salient color (e.g., red among green) which served as the critical distractor that captured attention. In addition, the distractor has a higher chance of appearing at a certain location (65%) compared to other locations (5% each). The result showed that, attentional capture was attenuated when the distractor appeared in the high probability location. In addition, reaction time for the target was slower when it appeared at the high probability location without the distractor once again confirming that processing of shapes appearing in the high probability location was suppressed. These results show that suppression can be applied to spatial locations based on a past *history* of distractors occurring there. For this reason, suppression is categorized as part of the history-driven mechanism (Theeuwes, 2018).

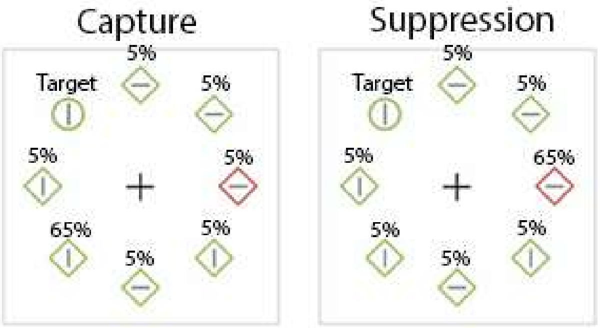


Figure 2. Schematic of a typical trial in Wang and Theeuwes (2018 a, b). A singleton distractor presented at a low probability location will capture attention (left) and a distractor presented at a high probability location will be suppressed (right).

1.4 History Driven Emotional Capture

Although there has been significant development in explaining how physically salient distractors capture attention or become suppressed, there is still much speculation about how emotional attention capture operates within the framework. According to the priority map hypothesis, emotional attention capture is categorized as history-driven capture because emotional capture essentially biases attention through *reward history* (Awh, Belopolsky, & Theeuwes, 2010). Studies show that a stimulus associated with reward or punishment (e.g., monetary, or electrical shocks) presented before the experiment or in the previous trial, still captures attention when presented as a totally task-irrelevant distractor in the subsequent trials (Anderson, Laurent, & Yantis, 2011a, b; Hickey, Chelazzi, & Theeuwes, 2010). An extreme real-world example would be PTSD patients who often experience involuntary flashbacks of past trauma. Although reward history is often discussed in terms of associations with monetary reinforcement or punishment by electrical shocks, it likely reflects a general emotional-based learning mechanism such as fear conditioning towards threatening animals, emotional facial expressions, and other types of emotional stimuli. The classical case of little Albert (Watson, & Rayner, 1920) demonstrated that fear of white rats (and similar white animals) can be learned by associating them with an adverse experience such as a loud sound. Also experiences of sickness can be associated very easily with a particular taste (taste aversion; Garcia, Kimeldorf, & Koeling, 1955).

It rather easy to label emotional capture by conditioning as history-driven capture because associations are arbitrary and experimentally induced. One can be sure that without conditioning, the neutral stimulus does not have the ability to capture attention. However, it is rather difficult to categorize attention capture by emotional images as history driven capture because it is difficult to quantify whether these images have been encountered before. Instead, it is entirely possible that the physical or perceptual properties of the emotional stimulus generates an automatic capture signal similar to physically salient stimuli.

1.5 Bottom-up and Top-down Emotional Capture

Is processing of emotional pictures stimulus-driven or goal-driven rather than being history driven? Similar to the early debate on physical salience, there has been a competition between the view of emotional capture as stimulus-driven versus top-down, goal-driven. The bottom-up accounts suggest that emotional attention capture is *automatically* driven by physical appearances (not referring to salient physical features) that are emotionally salient (Bannerman et al., 2012; Öhman et al, 2001). For example, Öhman and colleagues (2001) proposed that capture by a class of *biologically prepared* (Seligman, 1971) stimuli such as snakes, spiders, or emotional faces can occur extremely efficiently during visual search. Biological preparedness refers to the idea that humans have developed an innate biological inclination to process certain types of stimuli very efficiently for survival. So, when we encounter a snake, the representation of the snake is pre-attentively 'tagged' for selection and rapidly enters our post-attentive system.

In contrast, the top-down accounts suggest that all emotional information is first processed post-attentively. Once the emotional information is deemed motivationally salient, it rapidly elicits the necessary emotional responses (Lazarus, 1982). One could argue against the top-down view by pointing out that emotional distractors capture attention even when they are not relevant to the primary goal. However, the top-down view also claims that humans have a 'default mode' attentional set for processing emotional stimuli because it is always necessary to detect them to avoid harm (Öhman et al., 2001).

The bottom-up and top-down views also propose different neural pathways involved in emotional capture. Visual information travels along optic nerve fibers that connect the retina to the thalamus. Information is then relayed to the primary visual cortex and extra-striate visual areas for extensive processing. For top-down accounts, the amygdala would then receive information via feedback connections from various areas of the visual cortex. The amygdala is an almond shaped subcortical brain structure that is critically important in processing emotional stimuli (Whalen, 1998; Ledoux, 2000). Evidence suggests that a wide range of visual areas from V1 to later dorsal and ventral processing streams are modulated by top-down attention (Chelazzi, Miller, Duncan, & Desimone, 1993; Li, Piëch, & Gilbert, 2008; Nienborg, & Cumming, 2009; Womelsdorf, Anton-Erxleben, & Treue, 2008), which suggests that emotional attention capture can be modulated by attention as well.

In contrast, the bottom-up accounts suggest that visual information is conveyed through a subcortical pathway called the ‘low road’ that connects the retina to the amygdala, as opposed to the ‘high road’ that passes through the visual cortex (Ledoux, 1998, 2000; see Figure 3). Existence of the low road in humans is supported by the observation that patients with cortical blindness continue to process emotional stimuli (Morris, De Gelder, Weiskrantz, & Dolan, 2001; Pegna, Khateb, Lazeyras, & Seghier, 2005). In addition, diffusion tensor imaging shows that there is a neural pathway connecting the subcortical structures and the amygdala (Koller, Rafal, Platt, & Mitchell, 2019; Rafal et al., 2015). Recent studies show that activation in the amygdala can occur as early as 72 ms after onset of fearful faces, which is faster than responses in the fusiform face area (Méndez-Bértolo et al., 2016; Luo, Holroyd, Jones, Hendler, & Blair, 2007) which supports the existence of the low road.

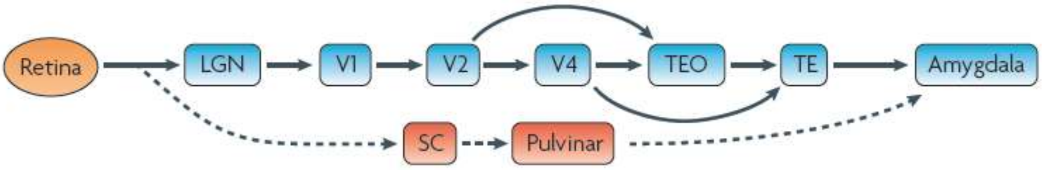


Figure 3. Schematic of visual processing of emotion. The typical emotional processing route (high road) is colored blue. Rapid emotional processing can occur through the subcortical route (low road) colored in red. LGN: lateral geniculate nucleus; TEO, TE: inferior temporal gyrus; SC: superior colliculus.

1.6 Criteria for Automaticity

The key to determining whether emotional picture processing is stimulus-driven or goal-driven relies on revealing whether or not emotional capture occurs automatically. Although this requires a precise understanding of what defines an automatic process, historically there has been little agreement on the features that define an automatic processes. The shortest list was provided by Yantis and Jonides (1990) that contained two features while the longest list was provided by Schneider, Dumais, & Shiffrin, (1984) that contained twelve (also see Hasher & Zacks, 1984; Posner & Snyder, 1975). Often the important features were different for different mental processes. For example, *speed* was considered important for automated acquired skills such as reading (Logan, 1997), because the execution of such skills become faster with practice.

Moore (2016) in a recent review acknowledges four features of an automatic process: Un/controlled, Un/conscious, Non/efficient, Fast/slow. In other words, an automatic process must not be controlled, can occur unconsciously, is efficient and fast. But how can we determine if a process does contain such features? For the purpose of this paper, we want to highlight two criteria that are frequently used in the attention capture literature. First, an automatic process is insensitive to increased perceptual or cognitive load of a concurrent information (Yantis & Jonides, 1990). Load is relevant to the Un/conscious feature of automaticity because increasing load also increases the threshold for consciousness towards a stimulus when the attentional capacity is consumed by a secondary task (Macdonald & Lavie, 2011). Load is also related to the Non/efficient feature because an automatic process that is efficient is likely to perform equally well under low attentional availability compared to when attentional resources are abundant.

Typically, perceptual load is increased by increasing the amount of stimulus information that is relevant to the task (Macdonald & Lavie, 2011, p1780). For example, detecting an X among the letter string ‘OOOOXOOOO’ s is easy because of the display contains only two types of information (the X and the O) thus perceptual load is low. However, detecting an X among a strong of random alphabet letters, such as ‘LWITYHSXLJWOHER’, increases perceptual load is high because now the amount of information is increased by the number of unique letters in the letter string. Cognitive load is typically increased by adding more information to be processed in working memory. This may include adding more items to memorize in a memorization task or increasing the digits of numbers in a mental calculation task.

The second criterion states that an automatic mental process is triggered by the presentation of a stimulus even when the participant does not *intend* to pay attention to that stimulus. In addition, an automatic process may not be countered by *goals* (stoppable, prevented, or interfered) once it has been triggered (Kahneman & Treisman, 1984; Moor, 2016; Yantis & Jonides, 1990). This is related to the Un/controlled feature of an automatic processes. This criterion is particularly relevant to capture and suppression. In the priority map framework, physically salient stimuli are deemed to generate a priority signal (or attend-to-me signal) which would capture attention without the intervention from goal-driven and history-driven mechanisms (Luck et al., 2021; Theeuwes, 2019). The priority signals itself, as opposed to the capture by the signal, is considered automatic and occurs in a purely stimulus-driven fashion. However, actual capture of attention can be terminated through reactive suppression (suppression in response to the stimulus) or even prevented through proactive suppression (suppression before the stimulus has been presented). While this seems like an indication that capture is not automatic, we still need to consider one more factor which is whether *goals* are necessary for suppression. Some claim that suppression (especially suppression towards features) is purely a result of selection history and is induced by repeated learning over several trials. Notably, instructing what to suppress trial-by-trial (i.e., manipulating explicit goals) is ineffective (Cunningham & Egeth, 2016; deVries et al., 2019; Gaspelin et al., 2019). Others suggest that trial-by-trial manipulations do induce suppression (Lien et al., 2010; Moher et al., 2011). However, even in cases that suggest trial-by-trial manipulations are ineffective at inducing suppression, explicit goals of the task are still known in advance and is not completely irrelevant to the task. This suggests that "some form of *deliberate* preparation is required to instantiate the appropriate control settings…" (Luck et al., 2021), that may be necessary for suppression. Another reason why suppression is may be highly relevant to the current goals is because, there is a debate regarding the priority map framework on whether the goal-driven and history-driven together form a single top-down *control settings* or control state. This has been a long stance of the top-down contingent theory which states, "These control settings, in turn, are a function of current behavioral, goals, as well as past experience or enduring biases of the organism." (Folk et al., 1992, p. 1043). For these reasons, the current paper considers that suppression is highly interrelated with the intention of the observer and that suppression of attention capture as an indicator that the capture process is not automatic.

It is important to note that there are instances in which researchers state that capture occurs automatically while paradoxically acknowledging that capture can be subject to control. For example, Luck et al., (2021) highlights one of the key characteristics of the new priority map framework by stating, "Certain kinds of stimuli … *automatically* generate a priority signal that, in the absence of specific attention control settings, will *automatically* capture attention." Here, the first use of the word automatic truly refers to the automatic nature of the priority signals. The second use however is more synonymous with “involuntary”, since the next sentence states that, “The capture of attention by salient singleton stimuli can be *prevented* if the attentional control system is appropriately configured”. Fabio and Capri (2019) also use the term automaticity in a similar way. They state that “… attention capture by threatening stimuli may be *automatic* by default but can be either *decreased* or *enhanced* by endogenous attention processes…”. Here they acknowledge that capture can be decreased which is technically a contradiction to the statement that capture by threatening stimuli is automatic. In a different section of the paper, they also state that “… threatening stimuli can capture attention, but this attentional effect cannot be regarded as totally bottom-up or automatic”. It is evident from this statement that the use of automaticity in the former statement was used roughly synonymous to “involuntary”. In this paper, we use the term automatic to refer to the former notion of automaticity (i.e., cannot be suppressed) of a mental process and is not used synonymously with involuntary capture.

1.7 Automaticity of Emotional Capture

The current section reviews conflicting findings regarding the automaticity of emotional capture. Studies that initially claimed that emotional capture was automatic demonstrated that capture simply occurs *involuntarily* when attention is directed away from the distractor toward a primary task stimulus as opposed to going undetected. For example, Vuilleumier and colleagues (2001) used fMRI to measure activation of the amygdala in response to pairs of task-irrelevant emotional faces presented in a vertical or horizontal configuration. The task was to determine whether the two faces were identical or not. In addition to the pair of face pictures, each display also contained a pair of house pictures so, for example, if the horizontal pair consisted of faces, the vertical pair would consist of houses and vice versa. The pair of faces could contain fearful and neutral expressions. A cue was provided at the beginning of the trial to indicate which orientation (horizontal or vertical) of the picture pairs was to be attended. The results showed that when participants were directed to attend to the house pictures, fearful faces still activated the amygdala suggesting that processing of emotion is involuntary and thus automatic. Other studies in which emotional distractors are presented as uninformative cues preceding the presentation of an attentional probe show further evidence of emotional capture. Typically, responses to the probe are facilitated when it is presented at the same location as the emotional cue (Armony & Dolan, 2002; Mogg, & Bradley, 1999; Pourtois, Grandjean, Sander, & Vuilleumier, 2004), suggesting that attention has shifted towards the emotional picture.

Meanwhile, studies that support a non-automatic, top-down emotional capture mechanism show that emotional capture is highly influenced by the amount of attention directed towards the emotional distractor. In fact, numerous studies show that emotional capture is modulated by task-context, attentional control, and attentional load (see Mohanty, & Sussman, 2013 for review). For example, Pessoa and colleagues (2002) demonstrated that increasing the difficulty of the primary task prevents emotional capture and eliminates responses in the amygdala. They presented an image of a face at the screen center and two bars with random orientations in the periphery. While fixating on the faces, participants were endogenously cued to attend to the faces or bars in the periphery. When attending to faces, the task was to report the gender of the face. While attending to the bars, the task was to report whether the orientation was similar or not. The expression of the faces could be neutral, fearful, or happy. Results show that when attending to the faces, the fearful and happy faces elicited greater amygdala responses compared to neutral. However, when performing a perceptually difficult orientation matching judgement task, the amygdala response to all expression types was absent, suggesting emotional processing is susceptible to top-down modulation. They concluded that attentional resources are required to process emotional distractors. The critical aspect of this paper is that the bar task was substantially more difficult than the matching task in Vuilleumier et al. (2001). Pessoa et al., point out that the performance accuracy of the bar matching task in Pessoa et al. which was 64%, was lower than the accuracy of the house matching task in Vuilleumier et al. which was 86%.

Automaticity of emotional capture has also been extensively studied using an experimental paradigm called Emotion Induced Blindness (EIB: Most, Chun, Widders, & Zald, 2005). Most and colleagues demonstrated that presentation of emotional distractors produced a transient blindness to a subsequent target presented at the same location. Emotional and neutral distractor images were presented in a rapid serial visual presentation (RSVP) stream of landscape images at a rate of 10 Hz (see Figure 4). The task was to report the orientation of a 90 degrees rotated landscape image presented 2 or 8 items after the distractor (lags 2 and 8). Results indicated that distractor presentation decreased target performance at lag 2 compared to baseline at lag 8. Critically, accuracy was lower when the distractor was an emotionally negative image than a neutral one. This suggests emotional distractors induce a greater transient blindness towards the target indicative of attention capture.

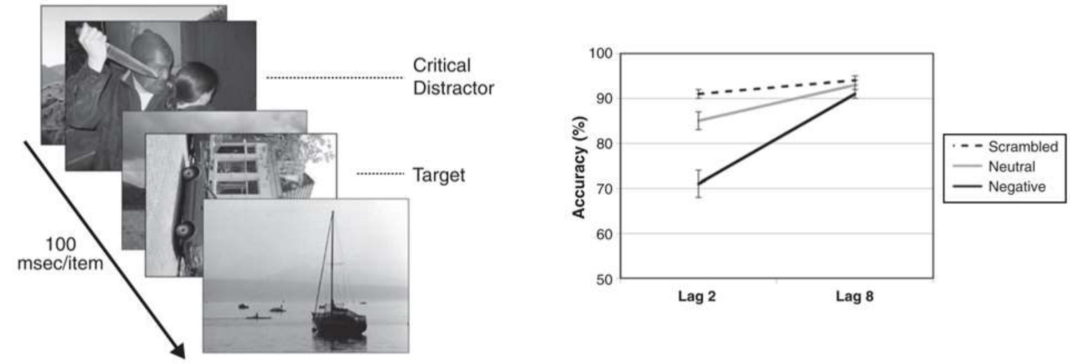


Figure 4. Schematic of the Emotion Induced Blindness procedure (left) and the results (right). Adopted from Most et al., (2005).

A follow up experiment tested whether EIB was automatic by altering the level of goal-driven attention and participant characteristics. The difference from the previous experiment was that the target image could be from either a specific category (e.g., natural landscape) or from a broader category (e.g., natural landscape and cityscape). Drawing the target from a more specific category would make the rejection of the distractor easier which would reduce EIB. The result show that EIB reduction only occurred in individuals with low Harm Avoidance (i.e., less susceptible to negative emotions). This suggests that emotional capture does not automatically occur and may be modulated by top-down attention, at least in individuals with low Harm Avoidance.

Baker, Kim, and Hoffman (2021) also tested automaticity of emotional capture in the EIB procedure by manipulating the physical salience of the emotional and neutral distractors. Physical salience is an important determinant of perceptual load and changing the relative saliency between stimuli can alter the level of interference caused by the distractors (Eltiti, Wallace, & Fox., 2005; Neokleous, Shimi, &

Avraamides, 2016). If emotional capture is diminished due to low distractor salience, this is an indication that emotional capture is susceptible to manipulations in perceptual load which violates the criteria for automaticity. Physical salience of the distractor was manipulated by presenting RSVP filler pictures that were physically similar or dissimilar to the distractors (see Figure 5). In a typical EIB study the differences in physical features between both distractors (emotional and neutral) and the filler pictures in the RSVP stream causes the distractors to perceptually 'pop-out' and direct bottom-up attention towards the distractor image. If EIB depends on physical salience for attention capture and this salience is reduced by using similar filler pictures in the RSVP stream, we would expect a corresponding reduction in the EIB effect. This would indicate that "emotional capture" depends on prior attention capture based on physical salience and that emotional salience does not contribute to the initial attention capture which is based only on physical salience. The results showed that EIB was reduced or eliminated when physical salience of the distractors was reduced, indicating emotional capture is not automatic.

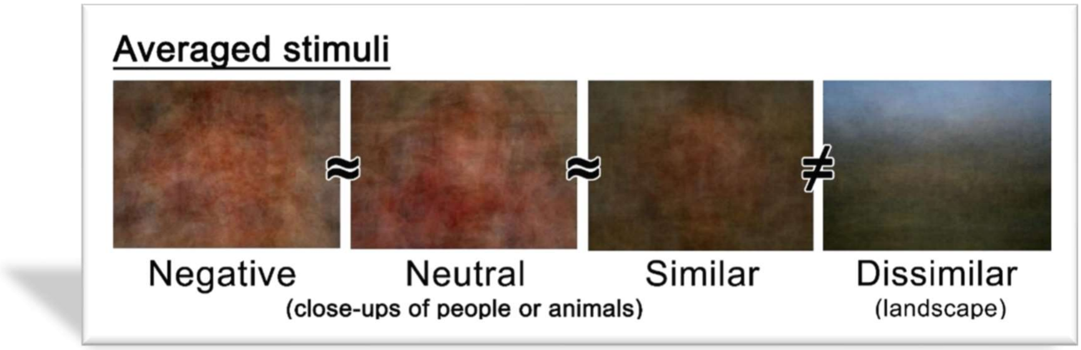


Figure 5. Aggregated images of each of the stimulus types. Negative, neutral and the similar stream images were matched in color and configuration while the dissimilar stream looks quite different from the rest.

Hoffman and colleagues (2020) tested automaticity of EIB by manipulating top-down goal driven attention in an ERP study. Goal driven attention was manipulated by adopting a difficult primary task that required directing attention away from the distractor. The primary task was multiple object tracking (MOT) which has been found to be so highly cognitively demanding that briefly presenting a salient distractor image in the background can go unnoticed due to inattention (Cohen, Alvarez, & Nakayama, 2011). During MOT, a small gap was presented on one of the tracked disks either simultaneously with or 300ms (0 & 300 ms SOA) after the distractor onset (see Figure 6). Participants were asked to make a speeded response to the orientation of the gap. If the distractors pictures in RSVP capture attention, as they do in EIB, it would delay responses to the gap when it occurs simultaneously with the distractor picture.

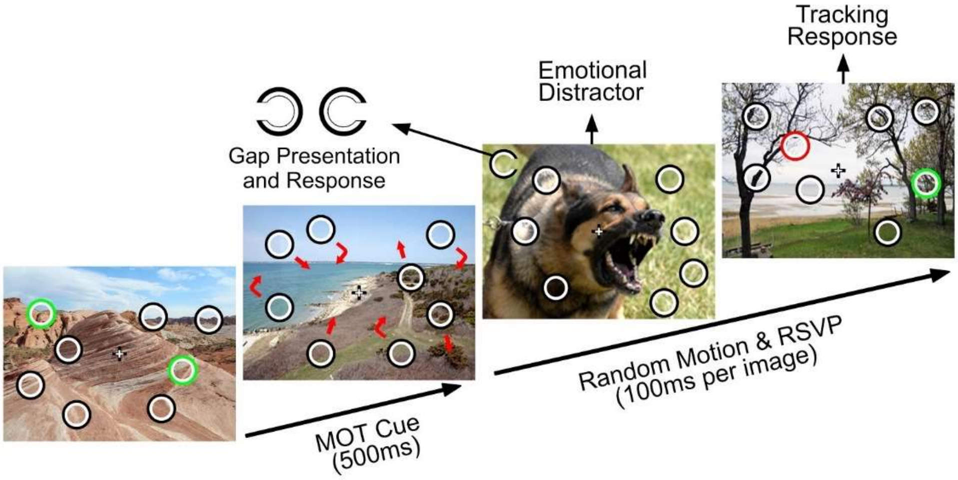


Figure 6. Schematic of the MOT task used in Hoffman et al. (2020) and Kim et al. (in prep). Target disks were tracked among other moving disks. A gap was presented for probing RT delays caused by the distractor presentation in Experiment 2 of Hoffman et al. and Kim et al.

The RT results showed that both emotional and neutral distractors delayed gap responses by 40 ms compared to a no-distractor baseline at 0 ms SOA, which suggests a common physical salience-based capture effect but not an emotional capture effect. Salience-based capture was gone by the time the gap was presented 300 ms after the distractor and also did not show an emotional effect. Interestingly, the emotional minus neutral subtraction of the ERP waveforms revealed a prominent Early Posterior Negativity (EPN; Schupp et al., 2006) component which is typically thought to reflect attention capture by emotional visual stimuli and occurs between 150 - 250 ms at the posterior electrode sites. The presence of the EPN without a behavioral emotional effect suggests that at some point the emotional valence of the pictures was extracted but without any evidence of attention capture by emotional pictures. This calls into question the claim that the EPN reflects emotional capture while the behavioral evidence suggests that emotional pictures do not invariably capture attention.

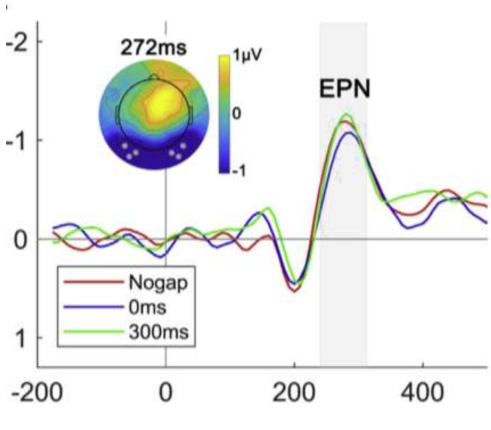


Figure 7. Graphs showing the EPN elicited by the distractors in Hoffman et al. Experiment 2. A significant EPN was generated even when RT did not show signs of emotional attention capture. Adopted from Hoffman et al., (2020).

Finally, Kim et al. (in prep) tested attention capture by emotional pictures and its dependence on physical salience while performing MOT. This study combined the manipulations used in Baker et al., and Hoffman et al. The task was identical to Hoffman et al., except that the background pictures in the RSVP were now either similar or dissimilar to the distractors as in Baker et al. We know that emotional effects on attention capture in EIB are largely eliminated when physical salience of the distractors is eliminated (Baker et al., 2021). If the EPN reflects greater attention capture by emotional pictures, it should also be eliminated in the similar background condition. Behavioral and ERP results in the dissimilar condition was consistent with Hoffman et al. There was the same 40 ms delay in detecting the gap in the presence of both emotional and neutral distractors at 0 ms SOA. And the emotional minus neutral ERP subtraction again showed a prominent EPN (see Figure 8; solid blue line). In contrast, the similar condition showed no delay in gap RT for either the neutral or emotional distractor at 0 ms SOA, consistent with the lack of EIB results of Baker et al. However, despite the lack of any evidence of emotional attention capture, the EPN was still present and comparable in magnitude to the EPN observed in the dissimilar condition in which behavioral evidence of attention capture was clearly present (see Figure 8; dotted orange line). These results are inconsistent with the claim that the EPN reflects attention capture.

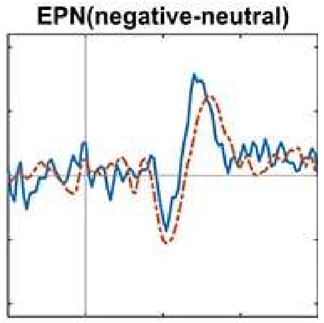


Figure 8. Graphs from Kim et al. (in prep). The negative minus neutral subtraction revealed a significant EPN both when the distractor was salient (dissimilar RSVP; solid blue line) and non-salient (similar RSVP; dotted orange line).

How can the EPN be an indicator while the gap RT does not show signs of capture while the EPN is present? This pattern of findings is not uncommon. Carretié (2014) reveals in a meta-analysis of emotional capture studies that many studies still report some form of capture related neural activity in the absence of behavioral indices. For example, an Emotional Stroop study found a prominent EPN but also showed no delays in target detection following emotional distractors compared to neutrals (Franken, Gootjes, & van Strien, 2009). A possible resolution may be that the EPN reflects an automatic component triggered by emotional stimuli that reflects a higher-level perceptual mechanism that processes emotional valence and calls for attention rather than reflecting the manifestation of emotional capture on the behavioral level. However, whether attention is captured may depend on the current top-down settings of the attentional system, which is similar to how physically salient distractors may capture attention.

1.8 Emotional Capture, Suppression and Spatial Attention

Baker et al., Hoffman et al. and Kim et al. focused on the temporal aspects of attention capture, rather than spatial shifts of attention, which were not required in these studies because the stimuli all occurred in the same location (at fixation). Currently, studies on the dynamics of spatial attention, emotional capture and suppression are scarce, and studying the interaction between these systems should provide insight on how the human visual attention mechanism operates and how it interacts with the emotional meaning of stimuli. One study used the additional singleton paradigm with classical conditioning (Kim, & Anderson, 2021) to study the spatial interaction of emotional capture and suppression. In this study, participants were trained to look for a shape singleton target among a homogenous display of distractor items. In the training phase, a physically salient color singleton appeared in one location with high probability and equal low probabilities in the remaining locations which induces active suppression via statistical learning (Wang & Theeuwes, 2018a, b). When distractors were absent, the target sometimes appeared in two different colors. One of the colors was associated with an emotional event such as monetary reward (Experiment 1) or mild electric shocks (Experiment 2). In the test phase, reward and punishment were no longer presented and the previous target colors were presented as singleton distractors that appeared with equal probability in all locations. The result revealed that singleton colors previously associated with reward or punishment yielded slower RTs compared to the non-associated color when presented in the low probability location. This confirms that *reward history* of the emotionally conditioned distractors can capture attention. In the high probability location, the RT for both conditioned and neutral distractors were faster compared to the low probability location which suggests that capture any type of distractor in the high probability location was suppressed. However, emotionally conditioned distractors captured attention more than neutral distractors in both high and low locations revealing an additive relationship between emotion and suppression. A strong indicator of automatic emotional capture would show that emotional distractors are not suppressed at all in the high location compared to the low location. However, this would not yet mean emotional capture is not automatic. It is important to note that the emotional singleton still captured more attention compared to neutrals in the high location. This suggests that suppression was not able to completely suppress emotional capture at the high location.

One notable feature of Kim, & Anderson is the use of classical conditioning in order to induce emotional attention capture (or Value-Driven Capture; Anderson, Laurent, & Yantis, 2011b). Emotional capture due to pure association as pointed out earlier can be categorized as history-driven capture. It is unclear whether their results would generalize to a set of emotional *images* that may be partially driven in a bottom-up fashion by the physical properties that bear emotional meaning.

1.9 Current Study

The main focus of the current proposal is to test whether capture of attention by emotional pictures still occurs in the face of active suppression, as would be predicted by the automaticity hypothesis. The current study addressed this issue using a variant of the *additional singleton paradigm* with statistical learning similar to Wang and Theeuwes (2018a, b), and Kim and Anderson (2021). Neutral or emotional singleton distractors were presented while participants searched for a shape singleton target. Probability of the singleton distractor appearance was increased at a particular location and this high probability location should be suppressed over repeated trials.

The claim is that the involuntary shift of spatial attention towards a salient object presented at the high probability location is prevented, meaning that capture is proactively suppressed. Unlike, Kim and Anderson who used classical conditioning for creating emotional distractors, the current study examines capture and suppression using a set of emotional and neutral *images* matched in physical salience.

Consistent with previous findings (Wang & Theeuwes, 2018a, b), we expected that neutral distractors would capture attention when presented at the low probability location and would be suppressed at the high location. Also, consistent with the claim that any stimulus presented at the high location is suppressed, even the target presented at this location would be initially suppressed until voluntary control over attention is regained. Also, one previous study suggests that emotional distractors at the high location are suppressed but still delayed RTs more than the neutral distractors (Kim & Anderson, 2021). This is seemingly a sign that emotional capture is not automatic because the emotional distractors were able to be suppressed, but it is still unclear why suppression was unable to completely eliminate the emotional interference effect at the high location. Perhaps while capture by physical salience is not automatic, capture by *emotional salience* is.

Here, it is important to acknowledge that, while neutral distractors are just physically salient, the emotional singletons are physically salient and also emotionally salient. This means that the emotional distractor has two independent factors potentially affecting attention capture. This is consistent with the priority map framework in the sense that stimulus-driven, goal-driven and the history-driven systems independently generates signals that call for attention which are later aggregated and projected onto the priority map. Because, physical salience and emotional salience are additive and independent, there is the possibility that only capture based on physical salience is suppressed and that capture by emotional salience may be immune to suppression, thus automatic. The reason why emotional salience might be automatic is because it is argued that emotional salience can be processed via a subcortical shortcut called the low road (LeDoux, 1998, 2000) that rapidly delivers sensory information to the amygdala for automatic processing of emotional information. If capture by *emotional salience* is automatic, the current study should initially replicate the findings from Kim & Anderson that show emotional interference effects in the high location with similar magnitude as the emotional interference at the low location.

Furthermore, Experiment 3 provided a more direct measure of spatial attention capture using mouse-tracking, rather than RT which provides a relatively indirect measure of spatial capture. Mouse tracking allows observations of shifts in spatial attention obtained with high temporal resolution recording of mouse movements during a reaching task. It is known that presenting a distractor that captures attention biases the mouse trajectory towards the distractor and the amount of deviation from the baseline indicates the strength of spatial attention capture (e.g., Dieciuc, Roque, & Boot, 2019). If capture by emotional salience is automatic while physical salience is not, the mouse trajectory for distractors appearing at the high location would show deviations towards the emotional distractor compared to the neutral distractor.

Chapter 2

EXPERIMENT 1

Experiment 1 uses the additional singleton procedure to test whether capture by emotional salience can be suppressed. Four objects were presented in an array in the shape of a cross (see Figure 9), one of which was a shape singleton target containing a bar that could be oriented vertically or horizontally. Participants had to report the orientation of the target bar while attempting to ignore a salient emotional or neutral singleton distractor. Singleton distractors appeared at the “high” location with a probability of 85%, while distractors appeared at the three low locations with a probability of which should induce suppression of any object presented at that location.

In the baseline condition, the singleton distractor was absent, and the target was presented at the low probability location. Capture was measured in terms of RT delays relative to the baseline when distractors and targets are both presented at low probability locations condition. Suppression was measured as the facilitation of RT in the high location compared to the low location. A second measure of suppression was indicated by the slowing in RT when the target appeared at the high probability location without a salient distractor in the display.

We anticipated several findings in this experiment. First, emotional, and neutral distractors should initially capture attention because they are both physically salient distractors presented among an array of homogenous set of stimuli. This would show as delayed RTs compared to the baseline when the singletons are presented in a low probability location. Second, there should be a facilitation in the RT for neutral singleton distractors presented in the high probability location compared to the low if suppression is successfully induced at the high location. A complete suppression effect would facilitate the RT in the high location to the point it matches the baseline, but we expect that there would be a partial suppression effect consistent with Wang and Theeuwes (2018a, b). Third, emotional singletons should involuntarily capture attention and cause additional interference compared to neutrals when they appear at the low probability location. A failure to observe emotional interference effect in the low location would indicate that emotional distractors do not always capture attention, given that the emotional distractor is sufficiently detectable by the visual system.

Perhaps the most obvious sign of automatic capture by emotional distractors would be if RT is not facilitated for emotional distractors at the high location compared to the low location. However, we expected that RT for emotional distractors at the high probability location should be faster compared to the low location consistent with a previous study (Kim & Anderson, 2021). Critically, if capture by *emotional salience* is automatic, we should be able to replicate the emotional interference effect at the high location that occurs with the same magnitude as the low location shown by Kim and Anderson.

2.1 Methods

2.1.1 Participants

A power analysis revealed that at least 10 participants were required to detect an effect size of 0.11 eta squared with 80% power which was the effect size of emotional attention capture reported in a similar recent study (Kim & Anderson, 2021). 60 participants were recruited on-line through Prolific and were compensated $5 (US) dollars for 30 minutes of participation. 11 participants were removed due to critical technical errors or overall response accuracy below 85%. Data from total of 49 participants (mean age: 22.2 / 36 male, 13 female) were included in the analyses. The participants were all right-handed, spoke English as their first or secondary language, and had normal or corrected-to-normal self-reported visual acuity. They were in college or had higher education degrees. None had a self-reported history of mental illnesses. Average trait anxiety (STAI; Spielberger, 1983) score was 45.49 with a standard deviation of 10.14.

2.1.2 Apparatus

The experiment was constructed using PsychoPy 3 which allows Python code to be translated into JavaScript that can be run on internet browsers. The experiment was uploaded onto the Pavlovia cloud server which provided hosting services for online distribution of Psychopy experiments and database management of the results. The participants used their own computer devices (desktop or laptop). To compensate for the difference in the individual monitor dimensions, the virtual chin rest technique (Li, Joo, Yeatman, & Reinecke, 2020) was used to determine the unknown monitor size and viewing distance for each observer in order to maintain a constant visual angle of the stimulus display across participants. The average monitor height was 20.8 centimeters (≈ 8.18 inches) and the average viewing distance was 68.54 centimeters (≈ 27.98 inches). The average frame rate of the monitors was 60.00 Hz. On average 1.22 centimeter was equivalent to 1 visual degree angle (vda).

2.1.3 Stimuli

All stimuli were presented against a grey background. Stimuli were comprised of singleton distractor pictures, filler distractor pictures and a target picture presented on a grey background. Singleton and filler pictures were square shaped while the target was in the shape of a circle (see Figure 9). The diameter of the target and the size of the distractors were 4° × 4°. On each trial, a target and three distractor items were displayed in a “+” formation and the total size of the stimulus display subtended 13° × 13°. A small fixation in the shape of a “+” (1° × 1°) appeared at the screen center. The center of each item was 4.5° from the center of the screen. Since the size of the stimuli were 4° × 4°, each stimulus fell within the parafoveal region 2.5° to 6.5° from the central fixation. The target and filler pictures contained high contrast gratings with a random orientation of 0° or 90°. At least one of the bar orientations differed from the rest within a trial. Singleton distractors were either emotionally negative (e.g., medical trauma, dangerous animals, threatening social situations) or neutral images (e.g., people and animals in an emotionally neutral context). The images were matched in terms of physical salience (spatial frequency, luminance, and color) using the natural images statistics toolbox (Bainbridge & Oliva, 2015; see Appendix A for matching results). The stimuli set was also roughly matched in terms of layout and the number of people or animals contained in the picture. The 70 negative and 70 neutral images were selected based on emotional valence and arousal ratings (Lang, & Bradley, 2007) collected prior to the experiment (see Appendix B for detail). Compared to neutral distractors (Valence ranging from 1 (positive) to 9 (negative): *M* = 4.201, *SD* = 0.544; Arousal ranging from 1 (low) to 9 (high): *M* = 4.034, *SD* = 0.390), the negative distractors (Valence: *M* = 7.654, *SD* = 0.614; Arousal: *M* = 6.716, *SD* = 0.328) were significantly more negative in terms of valence, $t(69) = 43.443$, $p < .001$, d = 5.192, and were higher in arousal $t(69) = 41.738$, $p < .001$, d = 4.989.

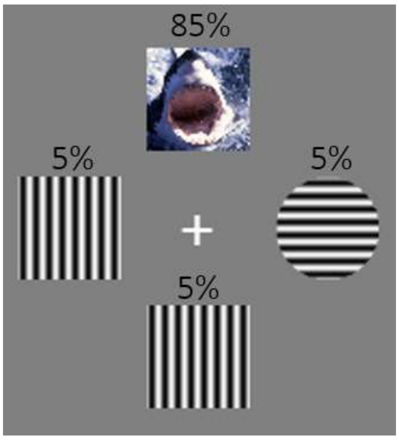


Figure 9. Schematic of a typical stimulus display in Experiment 1. Task is to report the orientation of the bars in a shape singleton (circle). The singleton distractor picture was either emotionally negative (i.e., negative valence, high arousal) or neutral (i.e., neutral valence, low arousal). When Distractors were present, there was a high chance of appearance at one of the four locations (the "High" location). There were equally low probabilities of appearance at each of the three "Low" locations.

2.1.4 Procedure

Online consent, the virtual chin rest procedure and the trait anxiety questionnaire was completed in that order, followed by the main instructions and 32 practice trials. In a typical trial, a fixation cross was presented for 1000 ms followed by the 4-item display that appeared for 3000 ms or until response. The task was to search for the circular target among 3 distractors. The target could occur in one of four possible locations (up, down, right, left) and participants had to report its orientation (0° or 90°) by pressing the 'z' or 'm' key. Participants were encouraged to make the response as accurately and as fast as possible. The response keys were counterbalanced across participants. Feedback regarding accuracy was displayed for 3000 ms when an inaccurate response was made. Feedback for response latency was also displayed for 3000 ms when the response was slower than 1500 ms. The two feedback displays could appear simultaneously.

The main trials contained 6 blocks of 92 trials each resulting in a total of 552 main trials and a rest period which was provided at the end of each block. 74% of the trials contained a singleton distractor and 26% of the trials contained three filler distractors and the target without the singleton distractor. In the main trials, one of the four stimulus location was randomly designated as the high probability location which was maintained throughout the experiment. When the singleton was present, there was an 85% chance that it would appear at the High location. Each of the three low locations had a 5% chance of a singleton appearing there. Negative and neutral valence pictures occurred equally often as singleton distractors. When the singleton was absent, the target appeared at each location with equal probability. When the singleton was present, the target always appeared at a low location, even when the singleton distractor did not occupy the high location. The high location was counterbalanced across participants. Participants were redirected to a completion URL once the experiment was completed.

2.1.5 Analysis

Prior to statistical analysis, data from block 1 was removed because repeated presentation of the distractor is necessary for inducing attention suppression (Wang & Theeuwes, 2018a, b; Luck et al., 2021). Because the main dependent measure was RT, individuals with an overall average error rate over 15% were removed from the analysis to minimize speed-accuracy tradeoff. For measuring RT, inaccurate responses, and responses faster than 200 ms were removed from the analysis.

Accuracy and RT data were analyzed using generalized linear mixed models (GLMM) using the glmer (lme4 package) function in R. GLMM has several advantages over ANOVA. First, GLMM can accommodate the differences in the distributions by using link functions that can be fitted for each individual data type. ANOVA assumes that the measurements form a normal distribution and violation of this assumption results in a reduction of statistical power. However, it is suggested that RT measurements have an inverse Gaussian distribution with a long tail reflecting a high concentration of RTs in the short latency range and fewer responses in the long RT range (Lo, & Andrews 2015; Stone, 2014). Proportion data such as accuracy with binary "correct" and "incorrect" responses form a binary logistic distribution with smaller variance as the data values occur at both of the extremes.

Another advantage of GLMM is that it can account for differences in variance introduced by an unbalanced number of observations between conditions. In the current experiment, there were more responses in the High condition compared to the Low condition. When there are more responses in the High conditions, the individual averages of the High conditions are less affected by outliers and the overall variance of the individual averages tend to be smaller than the Low condition. As a consequence, with traditional analyses, statistical power for detecting an emotional effect in the Low conditions would be considerably weaker compared to the High conditions.

2.2 Results

2.2.1 Percentage Error

Figure 10 shows the percentage error (PE) results for Experiment 1. Usually, attention capture would increase PE, according to Wang & Theeuwes, (2018a, b) and Kim & Anderson (2021), but presenting distractors at the low location did not increase error compared to the baseline while presenting the target at the high location did increase error. Also, they report the PE at the high location is lower than the low location but instead we found that when the distractors at the high location increased PE compared to the low location. Emotional distractors, did not increase PE compared to neutral distractors. It is unclear why the current PE results are different from previous reports. However, we focus on the RT measure since this is our main dependent variable.

The GLMM comparison between baseline (No-Dist-Targ-Low) and the rest of the conditions revealed a significantly lower PE in the baseline compared to the other conditions, $\chi 2(5, N = 49) = 28.691, p < .001$, w = 0.765. Tukey's post-hoc pairwise comparisons revealed that presenting the target at the High location (No-Dist-Targ-High) increased PE ($p < .05$). Presenting distractors (Neg-Low, Neut-Low, Neg-High, Neut-Low) overall did not increase PE compared to baseline ($p > .05$). A secondary two-way GLMM tested the distractor present conditions as a function of emotion and distractor location. No main effect of distractor valence was revealed, $\chi^2(1, N = 49) = 0.541, p = .462$, w = 0.105. The main effect of distractor location was significant $\chi^2(1, N = 49) = 4.290, p = .038$, w = 0.296. The two-way interaction was not significant, $\chi^2(1, N = 49) = 0.010, p = .919$, w = 0.015.

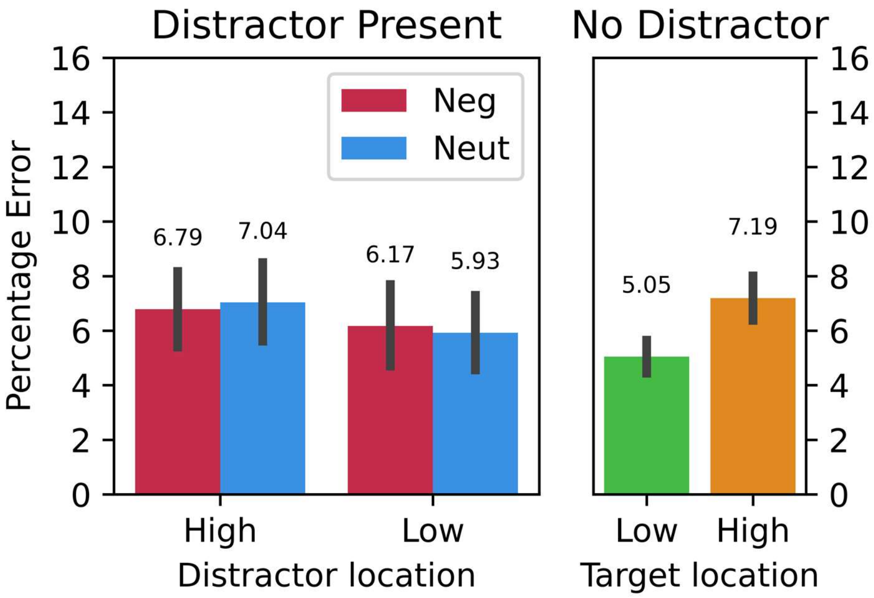


Figure 10. Error rate results from Experiment 1. Error bars indicate the SE. The green bar indicates the baseline (No-Dist-Targ-Low). The graph is based on the grand average and not on the model estimates.

2.2.2 Response Time

Figure 11 shows the response time (RT) results for Experiment 1. Compared to baseline, RT was slower when distractors were present (Neg-Low, Neut-Low, Neg-High, Neut-High) and when the target was presented at the High location without distractors (No-Dist-Targ-High). RT was faster when distractors appeared at the High location (Neg-High, Neut-High) compared to the Low location (Neg-Low, Neut-Low). RT was slower when Emotional distractors (Neg-High, Neg-Low) were present compared to Neutral distractors (Neut-High, Neut-Low).

The GLMM comparison between baseline (No-Dist-Targ-Low) and the rest of the conditions revealed a significant difference between all the conditions, $\chi^2(5, N = 49) = 97.42$, $p < .001$, w = 1.410. Tukey post-hoc tests showed that the RT in Neg-

Low, Neut-Low, Neg-High, and Neut-High were significantly slower than the baseline ($p < .05$). No-Dist-Targ-High was slower than baseline ($p < .05$).

A secondary two-way GLMM tested the distractor present conditions as a function of emotion and distractor location. A significant main effect of location was observed, $\chi^2(1, N = 49) = 24.008$, $p < .001$, w = .700, as well as a significant main effect of emotion, $\chi^2(1, N = 49) = 22.821$, $p < .001$, w = .682. The interaction was not significant $\chi^2(1\ N = 49) = 0.377$, $p = .539$, w = 0.088. Interestingly, pairwise comparisons revealed that Neg-High was slower than Neut-High ($p < .05$), but Neg-Low was not slower than Neut-Low ($p > .05$).

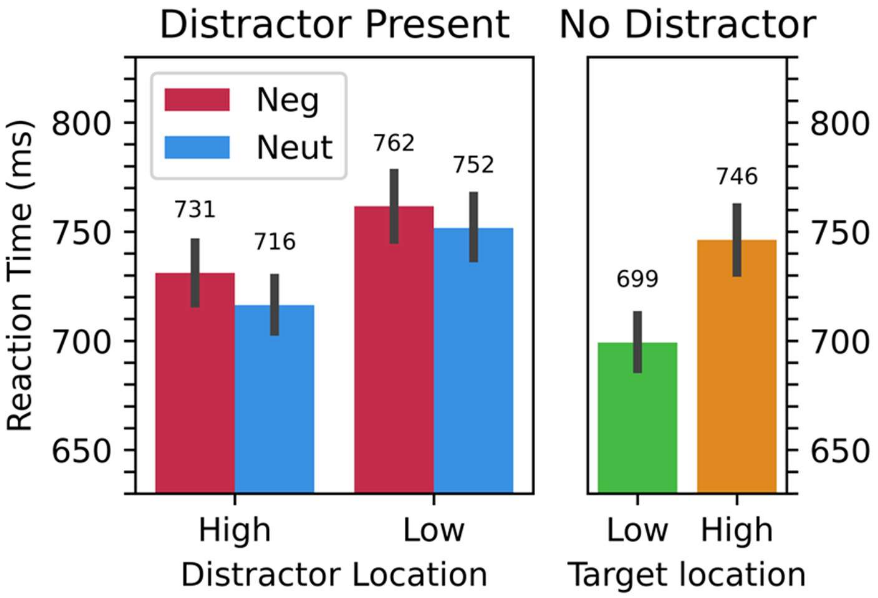


Figure 11. RT results from Experiment 1. Error bars indicate the SE. The green bar (No-Dist-Targ-Low) indicates the baseline. The graph is based on the grand average and not based on the modeling estimation.

2.3 Discussion

PE did not show clear signs of attention capture by distractors or suppression of the distractors. We did not observe an emotional interference effect as well. The only interpretable effect was the suppression effect towards targets presented at the high location when the distractors were not present. This indicates increasing the distractor appearance probability at the location suppresses processing of items at the location.

There was a clear sign of physical salience-driven attention capture shown by the RT delay when neutral distractors were presented in the low probability location compared to the baseline (see Figure 11). The neutral distractors presented in the high location show facilitations in RT compared to the low location, which is consistent with previous findings that physically salient distractors can be suppressed (Wang & Theeuwes, 2018a, b). Again, suppression was confirmed by the delayed RT for the target presented at the high probability location compared to the baseline.

Our main goal was to test whether emotional capture occurs automatically and our main criteria for determining automaticity was whether emotional capture can be suppressed at the high location after emotional capture involuntarily occurs in the low location. But critically, we did not anticipate the emotional distractors would be completely immune to suppression. Instead, we hypothesized that emotional distractors possess physical salience and emotional salience that are two properties which independently generate signals that calls for attention. It is known that capture by physical salience can be proactively suppressed but Kim & Anderson (2021) seem to suggest that capture by emotional salience is automatic. Now, if capture by *emotional salience* is automatic, then we should observe that emotional distractors should still produce emotional interference effects at the high location, even if emotional distractors were suppressed compared to the low location.

In short, we found evidence that emotional distractors were suppressed but captured more attention compared to neutrals (15 ms delay for emotional distractors) at the suppressed high probability location. This is consistent with the prediction that capture by emotional salience might be automatic but not capture by physical salience. However, we must approach this conclusion with caution. Prior to suppression, we found mixed results on whether emotional distractors captured attention involuntarily at the low location. The main effect of distractor valence and the lack of interaction between distractor valence and distractor location suggests that there was emotional capture at the low location. However, the post-hoc result suggests that there were no emotional interference effects at the low location. Therefore, it is ambiguous whether emotional capture occurs involuntarily even before testing for automaticity under suppression.

A factor that may have weakened emotional capture at the low location is the size of the stimuli. If the distractors are not large enough for the observer to perceive their content, extracting the valence of the image would be difficult due to a lack of sensory information, rather than the non-automaticity of emotional capture. This is especially true if the small stimuli are presented within the parafoveal regions in which visual acuity is poor compared to the fovea. Although there is evidence that emotional pictures can be detected in the parafoveal region, these stimuli also tend to be larger. For example, Calvo and Lang (2005) showed evidence of parafoveal emotional scene processing when the emotional image subtended 35.8° × 26.9° degrees within the parafoveal region. This is approximately 60 times the area size stimuli in the current experiment which were 4° × 4° in size. If small size of the emotional distractor weakened the emotional capture effect at the low location, how can there be emotional capture at the high location? Although there is no clear answer, we speculate that emotional capture in the low location was weak and not completely absent. Due to small differences in magnitude, emotional interference effects in the high location may have been significant but not in the low location. The reason why the emotional interference effect in the low location is smaller is perhaps due to random noise. Repeating the experiment with larger distractors should unveil the emotional effects in the low condition.

Experiment 2 increased the stimulus size compared to Experiment 1 in an attempt increase the visibility of the emotional capture effect in the low location. With increased size, the emotional distractors may involuntarily capture attention in the low location. The predictions on automaticity of emotional salience are the same as Experiment 1. If capture by emotional salience is automatic, there should be a main effect of emotion and a lack of interaction between the distractor valence and the location of distractor presentation. This would indicate that attention capture by emotional salience that involuntarily occurs at the low location is not suppressed even when presented at the high probability location.

Chapter 3

EXPERIMENT 2

Experiment 2 repeated Experiment 1 with larger stimuli that are also presented closer to the fovea. This was an attempt to strengthen the sensory representation of the distractors so that emotional information can be effectively analyzed by the perceptual system. We expect that increasing the stimulus size should ”uncover” the emotional interference effect in the low location that may have been too weak to reach significance in experiment 1.

In Experiment 1, the stimulus objects were presented in a “+” configuration which considerably limits their size and closeness to the screen center. In this configuration, increasing the size of the pictures would cause the pictures to overlap. A simple solution is to present the stimulus items in an “×” configuration (see Figure 15). This allows the stimuli to be much closer to each other and to fixation and increasing the size of the stimuli does not cause an overlap when they are stretched outwards in reference to the fixation. The predictions are identical to Experiment 1. We expect to find capture by physical salience and emotional salience at the low location. There should also be suppression of distractors and targets at the high location. Critically, if emotional salience captures spatial attention automatically, there should still be an emotional interference effect at the high location.

3.1 Methods

3.1.1 Participants

60 participants (mean age: 21.9 / 33 male, 27 female) were recruited on-line through Prolific and were compensated 5 US dollars for 30 minutes of participation. The screening criteria were identical to Experiments 1. Average trait anxiety (STAI) score was 46.57 with a standard deviation of 9.66.

3.1.2 Stimuli

The stimuli were identical to Experiment 1 except they were larger and were arranged in an "×" shape configuration instead of the "+" (see Figure 1). Each picture subtended 6.25° × 6.25° which was a 244.14% increase in area compared to the items in Experiment 1. The center of each stimulus item was 4.77° away from the screen center. Despite the increased size, X shaped configuration allowed the stimuli to be placed much closer to the center compared to Experiment 1. The total size of the outer bounds of the display items subtended 13° × 13° which is identical to Experiment 1.

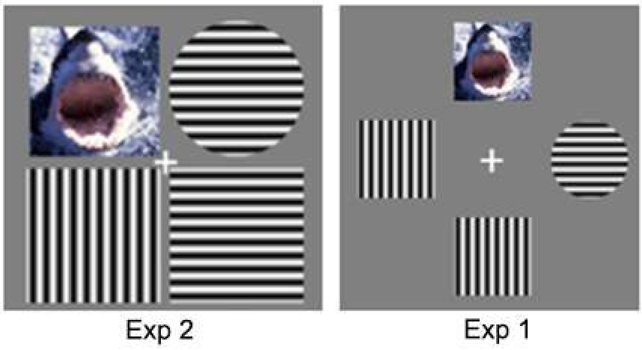


Figure 12. Schematic of a typical stimulus display in Experiment 2. The display items are similar to Experiment 1 except for stimulus size and configuration.

3.1.3 Procedure and Analysis

The procedure and analysis methods were identical to Experiment 1.

3.2 Results

3.2.1 Percentage Error

There was no significant effect of condition on Percentage Error (see Figure 13), $\chi^2(5, N = 60) = 9.863$, $p = .079$, w = 0.405.

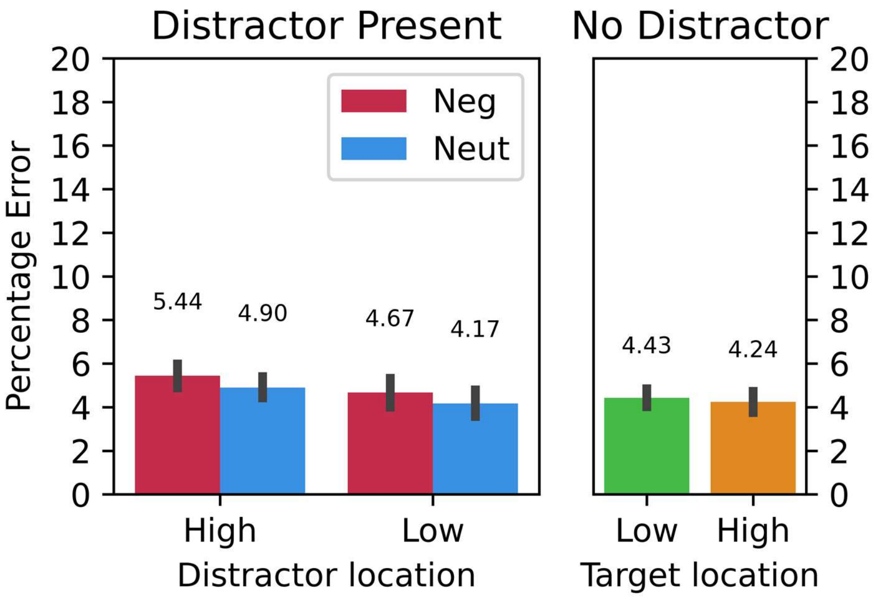


Figure 13. PE results from Experiment 2. Error bars indicate the SE. The green bar (No-Dist-Targ-Low) indicates the baseline. The graph is based on the grand average and not based on the results estimated from the model.

3.2.2 Response Time

Figure 14 shows that emotional distractors delayed RT by 46 ms, relative to baseline while RT for neutral distractor was not different from baseline. Absence of capture by neutral distractors was unexpected but the emotional distractor captured attention as expected. Compared to the low location, both emotional and neutral distractors presented at the high location showed faster RTs indicating suppression of the distractors. Target discrimination at the high location was also delayed compared to the low location, confirming that stimuli presented in the high location are suppressed. Most importantly, emotional distractors delayed RTs more than the neutral distractor at the high location. The magnitude of the emotional interference effect was statistically not different from the low location.

One-way GLMM revealed a significant difference between all six conditions, $\chi^2(5, N = 60) = 55.771, p < .001$, w = 0.964. Tukey's pairwise comparisons revealed that Neg-Low and Neg-High were slower than baseline ($p < .05$). No-Dist-Targ-High was also slower compared to baseline showing suppression ($p < .05$).

Two-way GLMM examining the factors of distractor valence and distractor location for the distractor present conditions showed a main effect of distractor location $\chi^2(1, N = 60) = 9.1685, p = .002$, w = 0.391. There was also a significant main effect of distractor valence $\chi^2(1, N = 60) = 14.715, p < .001$, w = 0.495. Emotional distractors delayed target detection compared to neutral distractors overall. The interaction was not significant, $\chi^2(1, N = 60) = 2.711, p = .100$, w = .213.

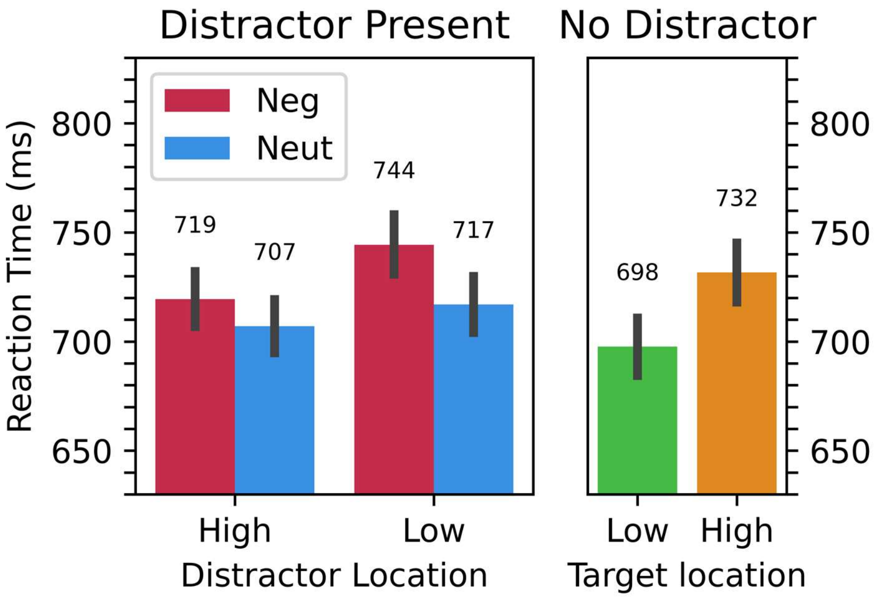


Figure 14. RT results from Experiment 2. Error bars indicate the SE. The green bar (No-Dist-Targ-Low) indicates the baseline. The graph is based on the grand average and not based on the results estimated by the model.

3.3 Discussion

Experiment 2 replicated Experiment 1 with a stimulus display using objects that were larger and closer to the fovea to increase the visibility of the singleton distractors. Unexpectedly, there were no differences in PE across conditions. The likely explanation is that target detection has become easier compared to experiment 1 due to the increase in the size of the stimuli and the reduction in distance from fixation. However, our main dependent variable was RT and inconsistencies in PE are less concerning.

RT results show that unlike Experiment 1, there was a significant emotional capture effect in RT at the low location indicating that emotional capture was triggered involuntarily by the emotional distractor images. Unexpectedly, there was a lack of capture by neutral distractors in the low location where physical salience was expected to result in attention capture. Perhaps the emotional capture effect in the low location was due to the absence of capture by neutral distractors. However, this does not undermine the fact that emotional distractors are able to capture attention involuntarily even when *equally* physically salient neutral distractors were not able to do so. In the high location, emotional and neutral distractors were both suppressed. Critically, the emotional distractors still delayed RT compared to neutrals in the suppressed high location. This might indicate while capture physical salience can be suppressed, capture by emotional salience can be automatic. If this is true, there should be a strong main effect of distractor valence and also an absence if interaction between distractor valence and location. An additional repeated measures Bayesian test between valence and location revealed, that there was indeed a very strong evidence for the distractor valence effect ($B_{01} = 0.002$) but only *anecdotal* evidence that the interaction supports the null hypothesis ($B_{01} = 1.360$). This is not a strong support that the emotional interference effect in the high location did not decrease due to suppression. More evidence is required to conclude that capture by emotional salience in the high probability location is actually automatic.

Another way of testing whether spatial attention is captured by emotional salience is using a more direct dependent variable that can show the changes in spatial attention with high temporal and spatial resolution. Note that presenting a salient distractor pulls *spatial* attention towards the distractor and the emotional interference effect in RT is supposed to reflect the time consumed by spatial attention to travel to the distractor, disengage and eventually reach the target (Theeuwes & Godijn, 2002). However, RT is a relatively *indirect* measure of spatial shift of attention. RT operates on the logical assumption that one could reveal the time taken by a certain mental stage by comparing two tasks that differed in just one factor (Donders, 1868). Therefore, we could compare the two RTs with emotional distractors and neutral distractors and infer that the time consumed to allocate spatial attention to the emotional distractor was greater than the neutral. However, the RT provides a measure of the sum of all stage durations preceding the button press rather than selectively reflecting the time for additional shifts of spatial attention. For this reason, researchers often later follow up with other methods such as eye tracking (e.g., Hoffman & Subramaniam, 1995), ERPs (e.g., Luck & Hillyard, 1994), or more recently mouse (or finger) tracking (e.g., Dieciuc et al., 2019; Moher, Anderson, & Song, 2015).

The advantages and disadvantages of these methods are clear. For example, ERPs are able to measure the neural processes involved in the capture process that occur prior to the response. However, it is challenging to link ERP components to underlying processes in a cognitive model of performance. For example, the EPN is often thought to reflect emotional capture but Hoffman et al., as well as Kim et al., found that the EPN can occur in the absence of behavioral measures of emotional capture suggesting that the EPN does not reflect capture per se. On the other hand, eye tracking and mouse tracking can measure spatial attention capture at the behavioral level. Eye tracking is extremely useful for measuring spatial attention because eye-movements are directly controlled via the frontal eye field areas which are also involved in programming covert spatial shifts of attention (Corbetta, & Shulman, 2002; Hoffman & Subramaniam, 1995; Thompson, Biscoe, & Sato, 2005). However, eye-trackers are not readily available for participants recruited on-line. Mouse-tracking (which is similar to finger or hand movement tracking) is a method that has been gaining popularity recently. According to previous studies on hand movements, the perceptual system evolved to enable organisms to conduct appropriate actions in the environment (Tipper, Lortie, & Baylis, 1992) and perceptual processing of the target and distractors during visual search independently create conflicting motor signals that at some point compete for action (Goldberg & Segraves, 1987). In other words, presentation of salient singleton distractors strongly influences our hand movements, in an almost reflexive way. It has also been shown that presentation of exogenous cues that matches the color of the target (i.e., contingent capture), which is known to shift spatial attention, also attracts the mouse-movement trajectories towards the salient cue, indicating that spatial attention and hand movements are highly inter-related (Dieciuc, et al., 2019; see intro from Experiment 3 for detailed description of the study). The biggest practical advantage of mouse-tracking over RT is perhaps that the spatial shift can be shown as coordinates in 2-d space with high temporal resolution. The mouse coordinates are sampled at similar refresh rates as a typical monitor which makes it convenient to record how spatial attention travels through the visual field.

For the purpose of this study, we focus on the mouse-tracking method due to its accessibility compared to other methods such as eye-tracking. A mouse is the most common input device for personal computers, unlike eye-trackers and EEG amplifiers that are still more accessible compared to other devices such as fMRI, but still much rarer compared to a mouse. Also, there is no requirement for setting up the device at the beginning of each session which allows the experiment to be administered over the internet without any assistance from the experimenter. This was extremely useful during the covid-19 crisis when in-person interaction with a participant was difficult.

Chapter 4

EXPERIMENT 3

There is evidence that attention capture and suppression can be reflected in our ongoing actions such as trajectories of our hand while performing a reaching task. (Boulenger et al., 2006; Dieciuc, et al., 2019; Finkbeiner, & Friedman 2011; Friedman, & Finkbeiner, 2010; Rheem, Verma, & Becker, 2018; Xiao, & Yamauchi, 2017). Unlike the RT method, which reflects the outcome of all stages leading to the response, mouse tracking can reflect real-time online processing across the spatial and temporal dimensions as attentional capture takes place (Dale, Kehoe, & Spivey, 2007; Freeman & Ambady, 2010; Kieslich & Henninger, 2017). Therefore, the mouse tracking method could show the direction of attention capture, suppression and whether attention capture occurred quickly or slowly (Dieciuc, et al., 2019).

For example, a recent study used mouse tracking to investigate the nature of attention capture using a version of the Posner cueing paradigm (Dieciuc, et al., 2019). In this study, 4 boxes containing Xs with 4 different colors were presented distributed equally on the upper part of the screen. The task was to selectively reach, using a mouse, for the target X defined by a specific color (e.g., Red). The target display was preceded by a non-informative, physically salient cue display either matching or not matching the color of the target. They found the mouse trajectory deviated most towards a matching target, but they also found a salient mismatching cue still captured attention and attracted the trajectory towards its location.

Moher, Anderson, and Song (2015) showed evidence of suppression using a finger reaching task which shares the same principal as mouse-tracking. The task was to make a reaching response with the finger towards a shape singleton target (i.e., a diamond among three squares). Occasionally one of the squares was presented in a salient color. The singleton distractor could either have high contrast (i.e., blue among red) or low contrast (i.e., pink among red). If high contrast items captured more attention, the reaching movement would deviate more towards the high contrast distractor, but instead, the results showed that the trajectory deviance was attenuated in response to the high contrast distractors compared to low contrast distractors. In a following experiment, Moher et al., observed similar results in value-driven capture. In this experiment, participants learned the association between colored circles and their monetary value by reaching for the target circles every trial. In a following test phase, these circles were presented as physically salient distractors (i.e., red, or green among grey items) while participants reached towards different grey scale stimuli. They found that deviations of the trajectories were *attenuated* by distractors associated with high value compared to the low value suggesting that the higher value distractors resulted in a greater suppression.

So far, there are no mouse-tracking studies that attempted to measure shifts of spatial attention in response to attention capture and suppression through statistical learning. However, Caro, Theeuwes, and Libera (2019) did this using eye-tracking which also measures the behavioral manifestation of attention capture with high temporal and spatial resolution, similar to mouse-tracking. Similar to the Wang and Theeuwes studies (2018a, b) they used the singleton paradigm in which the task was to report the orientation of a line contained in a grey circle among five other green circles that also contained randomly oriented lines. In spaces between the circles, a physically salient red circle (also containing a random oriented bar) was presented as a critical distractor. Two of the locations had a high probability of a distractor appearing there (High: 38%, Low: 6%) which should lead to suppression of the High locations. The participants responded with a keyboard and the eye movements were irrelevant to the task but the lines in the circles were so small that an eye-movement towards the target grey circle was necessary in order to make a response. The results showed that distractors presented in the low probability location had the lowest percentages of first saccades directed to the target compared to distractor absent conditions. The low probability trials also had the highest percentage of saccades directed to the distractors. This result is indicative of oculomotor capture by salient distractors that appeared in the low probability location. When the singleton distractor was presented in the high probability location, the percentage of trials that show first saccades to the target increased compared to when the distractor was presented at the low probability location. The percentage of trials with first saccades to the distractor decreased at the high location as well. This suggests that suppression of salient distractors induced through statistical learning is capable of *reducing* oculomotor capture that reflect shifts of spatial attention.

The purpose of Experiment 3 was to provide a direct observation of capture of spatial attention in response to emotional salience and whether it can be suppressed. Instead of a keyboard response, the task in Experiment 3 was to make a mouse-movement to a response box at a location corresponding to the target and make a mouse click according to the orientation of the bar. Other aspects of the study remained identical to Experiment 2.

First, we expected that when the singleton distractor is not present in the display and the target is at the low probability location, the mouse trajectory towards the target would be roughly a straight line connecting the initial starting point of the mouse (the fixation point)) and the target response box. Second, presenting a physically salient distractor at the low probability location would capture spatial attention and pull the mouse-trajectory towards the distractor. Emotional distractors are expected to capture more spatial attention than neutral distractors at the low location, as shown in Experiment 2. This means that the trajectory would deviate even more towards emotional distractors compared to neutrals. Third, the neutral singleton distractor presented in the high probability location should be suppressed and the deviation of the mouse trajectory should be attenuated compared to the neutral distractor appearing at the low probability location. We also predict that emotional distractors at the high location would be suppressed and attenuate the deviation of the mouse trajectory compared to the low location. Critically, we should consider the possibility that capture by emotional salience is automatic but not capture by physical salience. If capture by emotional salience is automatic, the mouse trajectory should show more deviation towards the emotional distractor compared to the neutral distractor at the high probability location.

4.1 Methods

4.1.1 Participants

120 participants were recruited on-line through Prolific and were compensated $5 (US) dollars for 35 minutes of participation. Number of participants were increased to ensure that sufficient high anxiety individuals would be included. 31 participants

were removed due to an average accuracy below 85%, excessive noise in the mouse trajectory data (determined through visual inspection) and screen refresh rate lower than 59Hz[1]. A Total of 89 participants (mean age: 21.9 / 61 male, 28 female) were included in the analysis. The screening criteria were identical to the earlier experiments. Average trait anxiety (STAI) score was 43.78 with a standard deviation of 9.65.

4.1.2 Apparatus

Participants with a desktop or laptop using a proper mouse (excluding trackpads or other similar cursor related input devices) were invited to enroll in the study. The virtual chinrest method was not used in this study because the entire stimulus display with the addition of a new response box needed to exactly fit the entire height of the screen. The experiment was conducted under the assumption that participants naturally maintain a reasonable viewing distance according to their screen size. Estimated viewing distances would approximate to 88, 74, 60 centimeters and one visual degree would approximate to 1.53, 1.29, 1.05 centimeters for 32-, 27-, and 22-inch (diagonal size) monitors, respectively.

4.1.3 Stimuli

The stimulus display was similar to Experiment 2 which produced a robust emotional capture effect. To accommodate mouse tracking responses, four white squares were added as responses boxes in the periphery (see Figure 15). The white

---

[1] The mouse trajectory is synchronized to the monitor refresh rate in this study independent from the mouse polling rate. Trajectory data would not contain sufficient detail with a low screen refresh rate.

boxes were also configured in an 'X' formation identical to the distractor and target items. The center of each response box was 15.56° away from the screen center. Size of the response box subtended 2.6° × 2.6°.

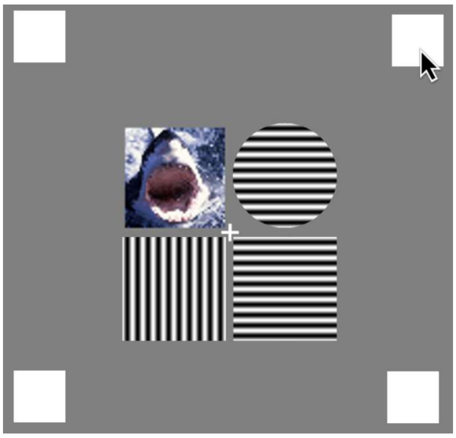


Figure 15. Schematic of typical display in Experiment 3. The central display items are identical to Experiment 2. Instead of keyboard responses, participants must judge the orientation of the target (circle), move the mouse from fixation to the white response box at the corresponding location as the target, and report the orientation by clicking the appropriate mouse button. The mouse movement must be initiated within 600 ms of stimulus onset.

4.1.4 Procedure

After providing consent online, participants filled out the Trait STAI questionnaire and were presented with the instructions. They were instructed to detect the circle target and report its bar orientation using a mouse. To carry out a correct response, the participants must move the mouse cursor to the response box appearing in the same direction as the target and click the left or right mouse button according to the orientation of the target. The assignment of mouse button to grating orientation was counterbalanced across participants. At the beginning of the trial, the 4 response boxes and the central fixation were presented without the distractor and target items. The trials were initiated only when the mouse cursor was placed directed at the fixation. This ensured that the cursor would be at the center of the screen at the beginning of each trial. The target and distractors appeared 1000 ms after trial initiation and were presented for 3000 ms or until response. Failure to move the mouse within 0.6 seconds or moving the mouse before the distractor and target presentation were scored as incorrect trials. The location of the High location was counterbalanced across subjects. The numbers of trials per condition was identical to the previous experiments. The practice contained 32 trials excluding the negative distractors with equal chance of distractors appearing at all locations.

4.1.5 Analysis

4.1.5.1 Mouse Trajectory Area

Mouse trajectory data was the main dependent variable of Experiment 3. We measured the area under the curve (AUC) of each trajectory from the distractor present conditions (Neg-High, Neg-Low, Neut-High, and Neut-Low) relative to the baseline trajectory (No-Dist-Targ-Low). The mean area of the trajectories was compared with a two-way repeated measures ANOVA with distractor valence and distractor location as factors.

In order to produce the trajectory area data, the mouse trajectory data were pre-processed with the following procedure which is similar to previous studies using mouse-tracking (e.g., Dale, Kehoe, & Spivey, 2007; Dieciuc, Roque, & Boot, 2019, Freeman & Ambady, 2010; Kieslich, Schoemann, Grage, Hepp, & Scherbaum, 2020). First, the number of the recorded coordinates for trajectories from each trial were matched using interpolation because the sampling rates between devices and the movement speed on each trial were different resulting in trajectories on each trial that contained variable numbers of digitized coordinates. In order to aggregate the trials, data for each trial was interpolated to contain 101 coordinates. After interpolation bad trials were removed through semi-automated visual inspection and were sorted into the following categories: type 1) mouse movement onsets before the distractor and target presentation, type 2) mouse movements being initiated 600 ms after distractor and target presentation, type 3) trials with distractors presented in the opposite direction of the target (because we want to measure deviation towards distractors presented adjacent to the target), type 4) trials with multiple mouse clicks, type 5) trials with random initial movements in the opposite direction from the target (because movement in the opposite X direction was meaningless and complicated the analysis), type 6) and extremely noisy movements detected through visual inspection. The cleaned data were re-arranged (using coordinate rotations and horizontal/vertical flips) so that the target position was at the upper-right corner and the distractor position was at the upper-left corner. The individual trials were aggregated and rotated 45° so that the straight trajectory from the center of the screen to the response box aligned with the x axis. This allowed measuring the deviations from the AUC with the values on the y-axis as the mouse traveled along the x-axis. So far, each data point in the trajectory was arranged in terms of equal latencies taken to reach the target. The datapoints were interpolated once more so that each data point represented an equal distance along the x axis. This step was necessary for measuring the AUC. Finally, jackknife re-sampling was performed in order to measure trajectory areas and reduce error variance (see Luck, 2014, p. 320). The area of the entire trajectory was measured by using the trapz function in Matlab for each condition from each individuals which was later grand averaged.

4.1.5.2 Mouse Click Percentage Error

Analysis of mouse click Percentage Error was similar to analyzing keyboard response Percentage Error in Experiments 1 and 2. The trials in which mouse movements were not initiate within 600 ms after stimulus onset were excluded from the trajectory area analysis but included in the mouse click PE analysis. The reason for inclusion in the PE analysis was because attention capture or suppression may have delayed the initiation of mouse movements causing response to be delayed beyond out 600 ms cutoff. For the mouse trajectory however, it was important that mouse movements were to be made within 600 ms of stimulus onset (see Kieslich et al., 2020).

4.1.5.3 Mouse Click RT

Mouse click RT was analyzed using GLMM similar to the RT data in Experiment 1 through 2.

4.2 Results

4.2.1 Mouse Trajectory

Figure 16 shows the grand average of the mouse trajectories in different conditions. Although it looks similar to an ERP waveform, the graph represents 2D coordinates rather than a time-series data. The x-axis represents the x-coordinates in 2D space which is also equivalent to the distance travelled from the fixation point, which was the starting point of the mouse movement, to the target response box where the response terminated with a mouse click. The value ‘1’ on the x-axis and the y-axis scale represents a normalized unit corresponding to the full height of the monitor. The Y axis represents the amount of deviation towards the singleton distractor using positive values and deviation towards the opposite direction from the singleton distractor with negative numbers. The x-axis represents the shortest distance travelled from the starting point towards the center of the target response box. As expected, in the baseline condition when the distractor was not presented and the target was presented at the low probability location (No-Dist-Targ-Low), (green line in Figure 16), the mouse travelled in a relatively straight line from the fixation towards the target response box.

The neutral distractor at the low location (blue solid line) captured attention and pulled the trajectory towards the distractor location. The negative distractor at the low location (red solid line) also pulled the trajectory towards it, also reflecting capture. However, the negative distractors pulled the trajectory more than the neutral singleton distractors towards the middle and end of the trajectories. This is confirmed by the fact that the area for negative distractors at the low location was larger than the neutrals (see Figure 17). This shows that emotional distractors captured additional attention involuntarily relative to neutrals.

At the high probability location, mouse trajectories for both emotional (red dotted line) and neutral (blue dotted line) distractors were pushed away from the distractor suggesting it was being suppressed. There was a difference in the size of the area between the negative and neutral at this location, but the difference was smaller compared to the low location and in fact the area for the negative distractor was greater than the neutral suggesting that the trajectory was further away from the distractor location. This suggests that spatial capture by emotional salience was *suppressed* when the distractor appeared at the high probability location.

One-sample t-tests reveal that the area in Neg-Low (*M* = 0.655, *SE* = 0.002), *t*(88) = 421.645, *p* < .001, d = 44.694, and Neut-Low conditions (*M* = 0.458, *SE* = 0.002), *t*(88) = 323.460, *p* < .001, d = 34.287, was significantly more positive than the baseline, suggesting that trajectories were pulled towards the distractors. Also, area in the Neg-High (*M* = -0.750, *SE* = 0.002), *t*(88) = -344.283, *p* < .001, d = -36.494, and the Neut-High (*M* = -0.737, *SE* = 0.002), *t*(88) = -371.541, *p* < .001, d = -39.383 were significantly more negative compared to the baseline suggesting they were pushed away from the distractors. Two-way rm ANOVA with distractor valence and distractor location as within-factors shows a significant main effect of distractor valence, *F*(1, 88) = 7031.040, *p* < .001, η_p^2 = 0.988, a significant main effect of distractor location *F*(1, 88) = 287718.948, *p* < .001, η_p^2 = 1.000, and a significant two-way interaction *F*(1, 88) = 11885.785, *p* < .001, η_p^2 = 0.993. Tukey's post-hoc comparison revealed that there were significant differences between negative and neutral distractor conditions in both locations (all *p*s < . 001). However, in the high location the trajectory area for the emotional distractor was further away from the distractor compared to the neutral distractor, which is opposite to the expectation that emotional distractors would pull the trajectory relative to the neutral. Therefore, the difference in the high location was not due to spatial capture of attention by emotional salience of the negative distractor.

An additional analysis of trajectory latency using the maximal deviation point shows that there was a significant difference in the latency at which the deviation reached its maximum *t*(87) = 7.374, *p* < .001. On average, Neut_Low was maximally deviant at 456 ms (*SE* = .0004) after the stimulus presentation while Neg_Low reached maximum deviation at 491 ms (*SE* = .005), showing a difference of 35 ms. This suggests emotional capture occurs later than physical salience-based capture.

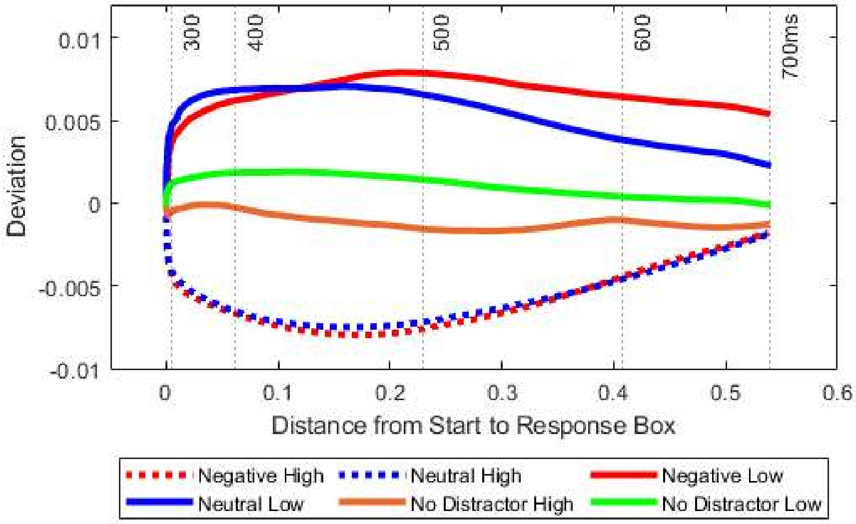


Figure 16. Grand Average of Mouse-Trajectory results from Experiment 3. The unit of the X and Y axis is normalized so that 1 refers to the full height of the monitor. X ranges from the starting point of the mouse (eye-fixation) and the final destination (response box). The Y axis represents deviation from 0. Positive Y values correspond to the direction of the distractors. The vertical dotted grey lines represent the averaged distance travelled at the labeled latency. The baseline (No-Dist-Targ-Low) is indicated by the green line.

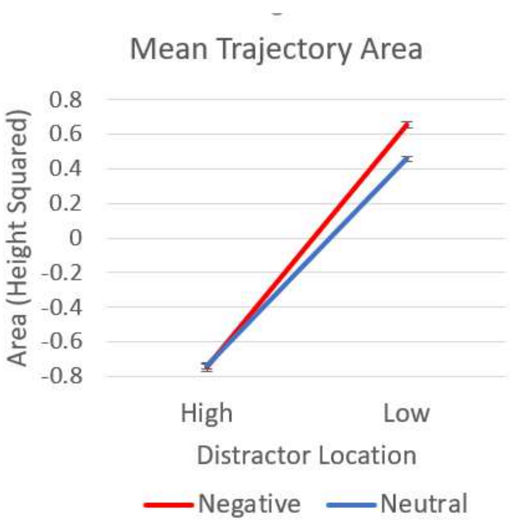


Figure 17. The grand averaged area under the curve of the trajectories when distractors were present as a function of distractor valence and location. The y-axis represents the size of the AUC. Although area is represented as a squared unit that cannot be negative, the polarity here represents the direction of the deviation (positive: towards distractor; negative: away from distractor). The error bars represent the SE which is extremely small due to reduced error variance from the Jackknife procedure.

4.2.2 Mouse Click Accuracy

Figure 18 shows PE across conditions. There were no significant differences across conditions, $\chi^2(5, N = 89) = 2.874$, $p = .720$, w = 0.180.

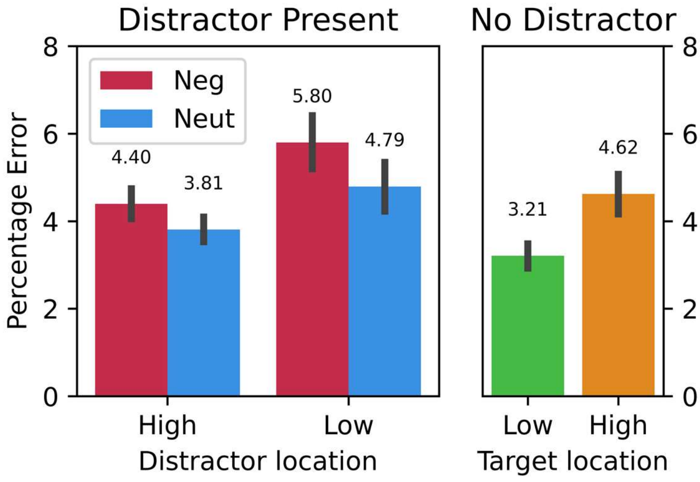


Figure 18. Mouse click error results from Experiment 3. Error bars indicate the SE. The green bar (No-Dist-Targ-Low) indicates the baseline. The graph is based on the grand average and not based on the results estimated by the model.

4.2.3 Mouse Click RT

Figure 19 shows the mouse click RT results for Experiment 3 which indicate that presenting neutral distractors at the low probability location delayed RTs compared to baseline reflecting attention capture by physical salience. Emotional distractors at the low location also delayed RTs compared to baseline and also compared to the neutral distractors suggesting that emotional capture occurred involuntarily. Compared to the low location, presenting neutral and emotional distractors at the high location facilitated RTs suggesting that these distractors were suppressed. Suppression was confirmed by the RT delay in detecting targets at the High (No-Dist-Targ-High) location. Even during suppression at the high location, the negative distractor delayed RT more than the neutral suggesting capture by emotional salience. There was no significant difference in the amount of emotional interference between the low and high locations. This is inconsistent with the trajectory area results that show a reduction of emotional interference effect in the high location compared to the low location.

One-way GLMM revealed that mouse click responses were different across the six conditions $\chi^2(5, N = 89) = 148.15, p < .001$, w = 1.286. Tukey's pairwise comparisons show that the baseline was faster than all other conditions ($p < .05$). Two-way GLMM revealed a significant main effects of distractor location, $\chi^2(1, N = 89) = 5.075, p = .024$, w = 0.239 and emotion, $\chi^2(1, N = 89) = 22.513, p < .001$, w = 0.503. There was no interaction of emotion and location. $\chi^2(1, N = 89) = 0.150, p = .698$, w = .041.

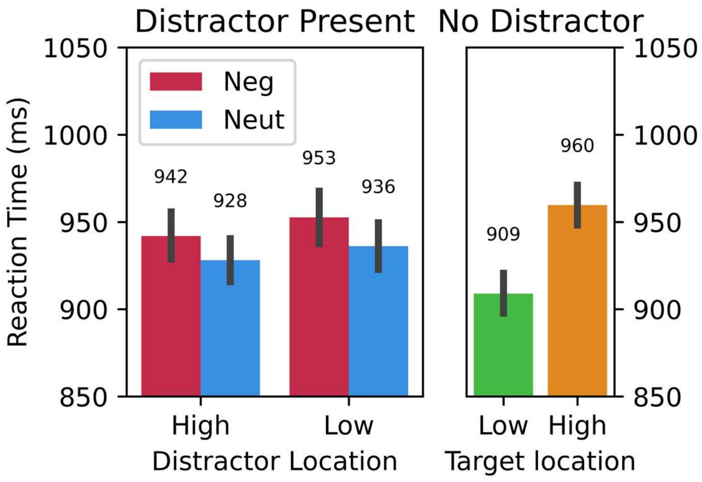


Figure 19. Mouse Click RT results from Experiment 3. Error bars indicate the SE. The green bar (No-Dist-Targ-Low) indicates the baseline. The graph is based on the grand average and not based on the results estimated by the model.

4.3 Discussion

Mouse click latency results replicated the main RT results from Experiments 2. Distractors at the low location captured attention and distractors at the high location were suppressed. Most importantly, emotional distractors still delayed RTs relative to neutrals and this emotional interference effect was not reduced or eliminated when they were suppressed at the high location.

Similar to mouse click RT, the mouse trajectories show evidence that emotional and neutral distractors at the low location *pull* the trajectory towards them suggesting capture of spatial attention. The deviation of the trajectory is also larger for emotional distractors at the low location suggesting that emotional capture occurred here. The latency of the trajectory was also 35 ms slower than the neutrals. This provides additional evidence that capture by emotional salience is slower compared to capture by physical salience. Note that automatic processes usually occur rapidly according to theories on automaticity (e.g., Moor, 2016), which may also suggest that emotional capture does not occur automatically (although speed is not the definitive criteria for automaticity in this study). Trajectories at the high location were *pushed* away from the distractor by emotional and neutral distractors. This means that spatial attention towards distractors presented at a suppressed location is not only reduced or prevented but is actively driven away from the distractor.

Most importantly, the emotional capture effect at the low location is absent at the high location. If emotional distractors capture spatial attention under suppression, the trajectory for emotional distractors should still be pulled more towards the distractor location compared to the neutral. But again, we do not find this shift in the trajectory. The lack of pull at the high location was not consistent with the click-RT data because the emotional interference effect was not eliminated in the high location according to the click-RT results.

Chapter 5

GENERAL DISCUSSION

The primary purpose of the current study was to test the automaticity of emotional attention capture under location-based suppression induced by statistical learning. Task-irrelevant emotional or neutral singleton distractor pictures were presented while participants engaged in a target orientation discrimination task across a series of three experiments. Location based suppression was induced by varying the probability of distractor presentation at a particular location (Wang, & Theeuwes, 2018 a, b). In the "High" probability condition, when distractors were present, there was an 85% chance that they appeared in a particular location. The remaining three "Low" locations each had a 5% chance of appearance. Distractors that appeared in the Low location should capture attention while distractors at the High location should be actively suppressed. The primary question was whether emotional distractors capture attention *automatically,* thereby evading suppression. According to the criteria for automaticity that suggests automatic processes should not be stopped or prevented (e.g., Kahneman & Treisman, 1984; Moor, 2016; Yantis & Jonides, 1990), we reasoned that if emotional capture was automatic, then capture by emotional distractors should not be suppressed when they are presented in the high probability location. We initially expected that signs of automatic emotional capture would show a pattern be that emotional distractors are completely immune to suppression at the high location. However, according to previous studies that measured RT as a dependent measure using a similar paradigm to ours (Kim & Anderson, 2021), emotional distractors were suppressed to some degree but not completely. There was a lingering emotional interference effect (i.e., the difference between emotional and neutral distractor RT) at high location that seemed to occur with similar magnitude to the emotional interference effect at the low location. Possibly this is because capture by physical salience can be suppressed but not capture by *emotional salience*.

Note that the emotional distractors in the current study are physically salient because they are pictures presented among homogenous grating objects. In addition, they are also emotionally salient, being more negative in valence and higher in arousal compared to the neutrals. According to the priority map framework, a distractor may produce signals that call for attention from three independent mechanisms: stimulus-driven, goal-driven, and history-driven. Because these signals stem from different sources rather than one unified origin, there is a possibility that suppression acts on these signals separately rather than together. In addition, protagonists of automatic emotional processing have argued that emotional information is rapidly and automatically conveyed to the amygdala through the subcortical ‘low road’ (Ledoux, 1998, 2000). If this is true, physically salient emotional distractors should be suppressed due to physical salience, but still capture more attention due to emotional salience compared to equally physically salient neutral distractors.

Overall, the RT results show that presenting a physically salient neutral distractor at a low probability location (Exp 1 & 3) captures attention compared to the no-distractor control. This reflects physical salience-based attention capture since the neutral distractor was a salient singleton image among homogenous stimuli items containing grating bars. In Experiment 2 and 3, emotional capture effects were observed in the low location showing that emotional distractors can capture attention involuntarily. Also, presenting distractors at the high location resulted in RTs that were consistently faster for both negative and neutral distractors compared to the low probability location which is consistent with the pattern of location-based suppression from previous studies (Wang & Theeuwes, 2018a, b; Kim & Anderson, 2021). This shows that even when emotional distractors capture attention involuntarily at the low location, they can be suppressed when presented at the high probability location.

Meanwhile, we also found that emotional distractors still delayed RT (Experiment 1, 2 & mouse-click RT from Experiment 3) more than neutrals at the high probability location. While this *might* be an indicator that capture by physical salience is not automatic while capture by *emotional salience* is, we find evidence in Experiment 3 with mouse-tracking that shows emotional interference with RT at the high probability location is not due to capture of spatial attention because the emotional effect in mouse-tracking disappears in the high location. Mouse tracking can reveal emotional capture and suppression effects on real-time shifts of spatial attention. This is because perceptual processing of the target and the distractors create conflicting motor signals that compete for action (Goldberg & Segraves, 1987). The procedure of Experiment 3 was identical to Experiment 2, but the keyboard button press was replaced by a combination of a mouse reaching movement and a decision expressed as a mouse click. Without any distractors in the baseline condition, the movement from the fixation to the target should roughly be a straight line, which was the case in Experiment 3. In the presence of attention capture, the trajectory of the movement should shift towards the distractor (e.g., Dieciuc, et al., 2019; Moher et al., 2015).

In the low location, Experiment 3 found that both emotional and neutral distractors pulled the trajectory towards the distractors reflecting spatial attention capture (see Figures 16 & 17) and this deviation was greater for emotional distractors compared to neutral distractors suggesting emotional spatial capture occurred involuntarily. In the high location, trajectories for both negative and neutral distractors were pushed away from the distractors suggesting that both distractors were suppressed. So far, the pattern of the results was similar to RT effects from Experiment 2 and Kim and Anderson (2021) that also show physical and emotional salience-based capture at the low location, and suppression at the high location.

Critically, we found evidence that emotional distractors, compared to neutral distractors, *do not* pull the trajectory towards the distractors in the high probability location. The results also show that the deviation due to emotional effects substantially decreased under suppression at the high location. This is an indicator that emotional salience does not capture spatial attention automatically. This is seemingly in contradiction to the RT results that indicated that emotional salience does capture attention under suppression. We suggest that presenting emotional distractors at a suppressed location does indeed prevent capture of spatial attention meaning that capture of spatial attention by emotional salience is *not automatic*, but it may take longer to suppress the spatial capture by emotional distractors compared to neutral distractors. To clarify, this does not mean that suppression is initiated later for emotional distractors compared to neutrals. Rather we suggest that the process of *preventing* or suppressing the shift of attention towards the emotional distractor takes longer than the neutral distractor even when they are equally physically salient, as they were in this study.

The current study supports previous studies that tested attention capture by emotional salience. Baker et al. (2021), reduced the physical salience of the emotional distractors embedded in an RSVP stream. The task was to detect a target that was preceded by an emotional distractor that typically interferes with target detection. The emotional distractor was physically salient when the surrounding RSVP images were perceptually distinct from the distractor, but physical salience was reduced when the RSVP images were perceptually similar to the emotional distractor. If physical salience of the emotional distractor is reduced, this leaves emotional salience as the only feature that can elicit attention capture. They found that reducing physical salience of the emotional distractor substantially reduced the emotional capture effect suggesting that emotional salience does not capture attention without accompanying physical salience. In other words, it is the physical salience of the emotional distractor which triggers capture of spatial attention and once the emotional distractor is attended, the emotional meaning of the picture is accessed resulting in additional interference, perhaps in the form of sustained attention to the emotional distractor.

However, there is still a possibility that some aspects of emotional capture are indeed automatic. Hoffman et al. (2019), showed that even when emotional distractors fail to capture attention, they produce an ERP component called early posterior negativity (EPN) that is isolated by subtracting the neutral distractor waveform from the negative distractor waveform. They tested whether emotional distractors in an RSVP stream captures attention during an MOT study. MOT is a cognitively demanding task known to cause temporary blindness towards stimuli at the background due to inattention (Cohen et al., 2011). The task was to track several disks that randomly moved in the screen and report a small gap that appeared in one the disks. They showed that the demanding MOT task successfully prevented emotional capture from taking place indicated by the RT results, but a prominent EPN component was still elicited.

Kim et al. (in prep), used the same MOT task to test if the EPN component occurs automatically due to purely emotional salience without physical salience. Physical salience was reduced with the method used in Baker et al. The results showed that without physical salience, the emotional distractor did not capture attention, as reflected in a lack of interference with a primary task. However, the emotional distractor still elicited a prominent EPN component that was similar to the EPN accompanying a physically salient emotional distractor that did show interference with the primary task. This suggests that while emotional salience without physical salience does not capture attention, emotional information is still processed as reflected in the EPN. Regarding the current study, it is possible that an automatic EPN component could be generated which could represent a just *call* for attention by emotional salience, even when the actual emotional salience-driven spatial capture in the high location is suppressed. Further work is required to confirm this hypothesis.

We have additional evidence that would support the non-automaticity of emotional capture. The mouse-trajectory data from Experiment 3 show that capture by emotional distractors in the low location produces a maximum deviation occurring 35 ms later compared to neutral distractors. If capture by emotional salience was automatic, emotional capture should also occur rapidly (Moor, 2016) at similar latencies compared to capture by neutral distractors. This suggests that emotional capture is delayed and possibly does not occur through the low road, but rather through the high road in which visual information travels through the primary visual cortex first before reaching the amygdala.

The current study also provides new evidence regarding the shift of spatial attention under location-based suppression. Previous studies suggested that suppression merely attenuates the deviations towards salient distractors (Moher, et al., 2015; Caro et al., 2019). Instead, we find that salient distractors completely reverse the trajectories in the opposite direction of the distractors. Moher, et al., (2016) suggested when a salient distractor is under suppression, the deviation of hand movement trajectory towards the distractor is attenuated. However, their study was not testing location-based suppression and is unclear whether this would generalize to the current studies. Caro and colleagues, (2019) used eye-tracking to observe how spatial-attention would be modulated under location-based suppression induced with statistical learning similar to the current study. They showed that when the distractors appeared at the high distractor location, the percentage of trials that show first saccades to the distractor decreased and the percentage of trials that show first saccades to the target increased compared to the baseline. This is a sign that the shift of spatial attention towards the distractor was reduced and was instead directed towards the target. This implies that the capture was attenuated but whether the trajectory of the eye saccade deviates towards the opposite location from the high distractor location is unknown. It is unclear why our study finds patterns that completely reverses the trajectory of the mouse-movement during suppression. Perhaps the discrepancies in the results arise due to differences in the probability manipulation or the layout of the stimulus. Further examination is required on this issue.

In summary, we provide crucial evidence in resolving whether capture of spatial attention by emotional salience is automatic. Emotional singleton distractors that would usually capture spatial attention involuntarily, can be prevented from capturing spatial attention when presented at a proactively suppressed location, suggesting that emotional salience-based attention capture is not automatic.

REFERENCES

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Learned value magnifies salience-based attentional capture. *PloS one*, *6*(11), e27926.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367-10371.

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia*, *40*(7), 817-826.

Arnell, K. M., Killman, K. V., & Fijavz, D. (2007). Blinded by emotion: target misses follow attention capture by arousing distractors in RSVP. *Emotion*, *7*(3), 465.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437-443.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485-496.

Bainbridge, W. A., & Oliva, A. (2015). A toolbox and sample object perception data for equalization of natural images. *Data in Brief*, *5*, 846-851.

Baker, A. L., Kim, M., & Hoffman, J. E. (2021). Searching for emotional salience. *Cognition, 214*, 104730.

Bannerman, R. L., Temminck, E. V., & Sahraie, A. (2012). Emotional stimuli capture spatial attention but do not modulate spatial memory. *Vision research*, *65*, 12-20.

Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of cognitive neuroscience, 18*(10), 1607-1615.

Broadbent, D. E. (1958) *Perception and communication*. New York: Oxford University Press.

Calvo, M. G., & Lang, P. J. (2005). Parafoveal semantic processing of emotional visual scenes. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 502.

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(4), 1228-1258.

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*(6427), 345-347.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological science*, *22*(9), 1165-1172.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience, 3*(3), 201-215.

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition, 35*, 15–28. http://dx.doi.org/10.3758/BF03195938

Di Caro, V., Theeuwes, J., & Della Libera, C. (2019). Suppression history of distractor location biases attentional and oculomotor control. *Visual Cognition, 27*(2), 142-157.

Dieciuc, M. A., Roque, N. A., & Boot, W. R. (2019). The spatial dynamics of mouse-tracking reveal that attention capture is stimulus-driven rather than contingent upon top-down goals. *Journal of Experimental Psychology: Human Perception and Performance, 45*(10), 1285.

Donders, F. C. (1868). Die schnelligkeit psychischer processe: Erster artikel. Archiv für Anatomie, *Physiologie und wissenschaftliche Medicin*, 657-681.

Eltiti, S., Wallace, D., & Fox, E. (2005). Selective target processing: Perceptual load or distractor salience?. *Perception & Psychophysics, 67*(5), 876-885.

Fabio, R. A., & Caprì, T. (2019). Automatic and controlled attentional capture by threatening stimuli. *Heliyon, 5*(5), e01752.

Finkbeiner, M., & Friedman, J. (2011). The flexibility of nonconsciously deployed cognitive processes: evidence from masked congruence priming. *PloS one, 6*(2), e17095.

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*(4), 1030.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology-General, 130*, 681-700.

Franken, I. H., Gootjes, L., & van Strien, J. W. (2009). Automatic processing of emotional words during an emotional Stroop task. *Neuroreport, 20*(8), 776-781.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods, 42*, 226–241. http://dx.doi.org/10.3758/BRM.42.1.226

Friedman, J., & Finkbeiner, M. (2010). Temporal dynamics of masked congruence priming: evidence from reaching trajectories.

Garcia, J., Kimeldorf, D. J., & Koelling, R. A. (1955). Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science, 122*, 157–158.

Geng, J. J. (2014). Attentional mechanisms of distractor suppression. *Current Directions in Psychological Science*, *23*(2), 147-153.

Glickman, M., & Lamy, D. (2018). Attentional capture by irrelevant emotional distractor faces is contingent on implicit attentional settings. *Cognition and Emotion, 32*(2), 303-314.

Goldberg, M. E., & Segraves, M. A. (1987). Visuospatial and motor attention in the monkey. *Neuropsychologia, 25*(1), 107-118.

Grose-Fifer, J., Rodrigues, A., Hoover, S., & Zottoli, T. (2013). Attentional capture by emotional faces in adolescence. *Advances in cognitive psychology*, *9*(2), 81.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39*(12), 1372–1388.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience, 30*(33), 11096-11103.

Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. *Journal of cognitive neuroscience*, *21*(4), 760-775.

Hinojosa, J. A., Mercado, F., Albert, J., Barjola, P., Peláez, I., Villalba-García, C., & Carretié, L. (2015). Neural correlates of an early attentional capture by positive distractor words. *Frontiers in psychology*, *6*, 24.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & psychophysics, 57*(6), 787-795.

Hoffman, J. E., Kim, M., Taylor, M., & Holiday, K. (2020). Emotional capture during emotion-induced blindness is not automatic. *Cortex*, *122*, 140-158.

Huang, Y. M., Baddeley, A., & Young, A. W. (2008). Attentional capture by emotional stimuli is modulated by semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 328.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automacy. In R. Parasuraman and D. A. Davies (Eds.), *Varieties of attention*. New York: Academic Press.

Kieslich, P. J., Schoemann, M., Grage, T., Hepp, J., & Scherbaum, S. (2020). Design factors in mouse-tracking: What makes a difference?. *Behavior Research Methods, 52*(1), 317-341.

Kim, H., & Anderson, B. A. (2021). Combined influence of valence and statistical learning on the control of attention: Evidence for independent sources of bias. *Cognition, 208*, 104554.

Kim, M., Aul, C., & Hoffman, J., (in preparation). Emotional Salience does not capture but automatically calls for attention.

Koller, K., Rafal, R. D., Platt, A., & Mitchell, N. D. (2019). Orienting toward threat: Contributions of a subcortical pathway transmitting retinal afferents to the amygdala via the superior colliculus and pulvinar. *Neuropsychologia*, *128*, 78-86.

Lang, P., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. *Handbook of emotion elicitation and assessment*, *29*, 70-73.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American psychologist*, *37*(9), 1019.

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual review of neuroscience*, *23*(1), 155-184.

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). controlling for participants' Viewing Distance in Large-Scale, psychophysical online experiments Using a Virtual chinrest. *Scientific Reports*, *10*(1), 1-11.

Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, *50*(6), 951-962.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyze reaction time data. *Frontiers in psychology*, *6*, 1171.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & writing quarterly: Overcoming learning difficulties, 13*(2), 123-146.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000.

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition, 29*(1), 1-21.

Luo, Q., Holroyd, T., Jones, M., Hendler, T., & Blair, J. (2007). Neural dynamics for facial threat processing as revealed by gamma band synchronization using MEG. *Neuroimage*, *34*(2), 839-847.

Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics, 73*(6), 1780-1789.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657-672.

Méndez-Bértolo, C., Moratti, S., Toledano, R., Lopez-Sosa, F., Martínez-Alvarez, R., Mah, Y. H., ... & Strange, B. A. (2016). A fast pathway for fear in human amygdala. *Nature neuroscience*, *19*(8), 1041-1049.

Mogg, K., & Bradley, B. P. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition & Emotion*, *13*(6), 713-740.

Mohanty, A., & Sussman, T. J. (2013). Top-down modulation of attention by emotion. *Frontiers in Human Neuroscience*, *7*, 102.

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.

Moher, J., Anderson, B. A., & Song, J.-H. (2015). Dissociable effects of salience on attention and goal-directed action. *Current Biology, 25*, 2040–2046.

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual review of psychology, 67*, 263-287.

Morris, J. S., DeGelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain, 124*(6), 1241-1252.

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*(4), 654-661.

Most, S. B., Smith, S. D., Cooter, A. B., Levy, B. N., & Zald, D. H. (2007). The naked truth: Positive, arousing distractors impair rapid target perception. *Cognition and Emotion*, *21*(5), 964-981.

Neokleous, K., Shimi, A., & Avraamides, M. N. (2016). Modeling the effects of perceptual load: Saliency, competitive interactions, and top-down biases. *Frontiers in psychology, 7*, 1.

Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature*, *459*(7243), 89-92.

Öhman, A. (2009). Of snakes and faces: An evolutionary perspective on the psychology of fear. *Scandinavian Journal of Psychology*, *50*(6), 543-552.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466.

Padmala, S., Sambuco, N., Codispoti, M., & Pessoa, L. (2018). Attentional capture by simultaneous pleasant and unpleasant emotional distractors. *Emotion*, *18*(8), 1189.

Pegna, A. J., Khateb, A., Lazeyras, F., & Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nature Neuroscience, 8*(1), 24-25.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, *51*(1), 195-203.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences*, *99*(17), 11458-11463.

Posner, M. I., & snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, NJ: Erlbaum.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex*, *14*(6), 619-633.

Rafal, R. D., Koller, K., Bultitude, J. H., Mullins, P., Ward, R., Mitchell, A. S., & Bell, A. H. (2015). Connectivity between the superior colliculus and the amygdala in humans and macaque monkeys: virtual dissection with probabilistic DTI tractography. *Journal of Neurophysiology*, *114*(3), 1947-1962.

Rheem H, Verma V, Becker DV. (2018) Use of Mouse-tracking Method to Measure Cognitive Load. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2018;62(1):1982-1986. doi:10.1177/1541931218621449

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics*, *72*(6), 1455-1470.

Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition and Emotion*, *29*(4), 687-694.

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 1-27). Orlando, FL: Academic Press.

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in brain research*, *156*, 31-51.

Seligman, M. E. (1971). Phobias and preparedness. *Behavior therapy, 2*(3), 307-320.

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.

Spielberger, C. D. (1983). Manual for the State-Trait Anxiety Inventory STAI. Palo Alto, CA: Mind Garden.

Stone, J. V. (2014). Using reaction times and binary responses to estimate psychophysical performance: an information theoretic analysis. *Frontiers in neuroscience*, *8*, 35.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599-606.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77-99.

Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition, 1*(1):29, 1-15.

Theeuwes, J. (2019). Goal-driven, stimulus-driven, and history-driven selection. *Current Opinion in Psychology*, *29*, 97-101.

Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception & psychophysics, 64*(5), 764-770.

Thompson, K. G., Biscoe, K. L., & Sato, T. R. (2005). Neuronal basis of covert spatial attention in the frontal eye field. *Journal of Neuroscience, 25*(41), 9479-9487.

Tipper, S. P., Lortie, C., & Baylis, G. C. (1992). Selective reaching: Evidence for action-centered attention. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 891–905.

Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin, 20*(1), 12-16.

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*(7), 2266-2279.

Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology*, *56*(2), 153-158.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*(3), 829-841.

Wang, B., & Theeuwes, J. (2018a). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*, *80*(4), 860-870.

Wang, B., & Theeuwes, J. (2018b). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 13.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of experimental psychology, 3*(1), 1.

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.

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current directions in psychological science*, *7*(6), 177-188.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1-8.

Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *Journal of Neuroscience*, *28*(36), 8934-8944.

Xiao, K., & Yamauchi, T. (2017). The role of attention in subliminal semantic processing: A mouse tracking study. *PloS one, 12*(6), e0178740.

Yamaguchi, M., & Harwood, S. (2015). The influence of fear-inducing stimuli on learning of visual context. *Perception, 44*, 28.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *Journal of Experimental Psychology: Human perception and performance*, *10*(5), 601.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance, 16*(1), 121.

Appendix A

CAPITALIZED APPENDIX TITLE

Table 1. Differences in spectral power of spatial frequency between sets of negative and neutral distractor images. The p-value is obtained through an independent t-test. There were no differences in spectral power overall and also in any spatial frequency band.

Spatial Frequency Band	**Low** 10%	30%	50%	70%	**High** 90%	**Total**
p-value	.238	.870	.691	.561	.376	.587

Table 2. Differences in color intensity between sets of negative and neutral distractor images in LAB color space. The p-value is obtained through an independent t-test. L shows differences in luminance. A (red-green axis) and B (yellow-blue axis) shows differences in hue. Results do not show differences in color intensity and luminance.

	L	**A**	**B**
p-value	.510	.884	.833

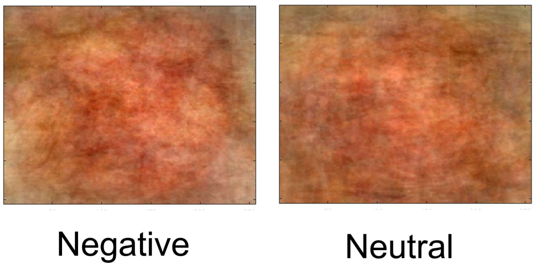


Figure 20. Aggregated images from all the distractors from the negative set and the neutral set. The two sets overall show prominent red across the screen with darker contours towards the middle.

Appendix B

DISTRACTOR EMOTIONAL RATINGS

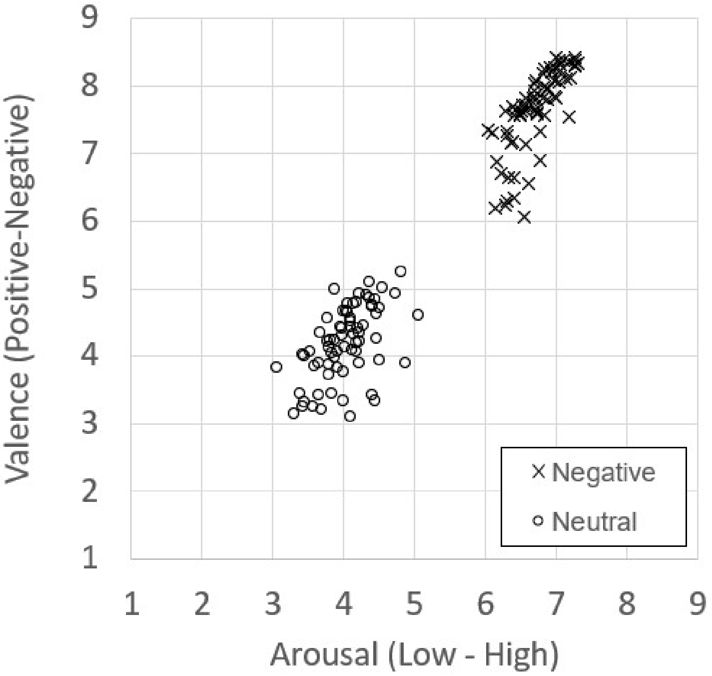
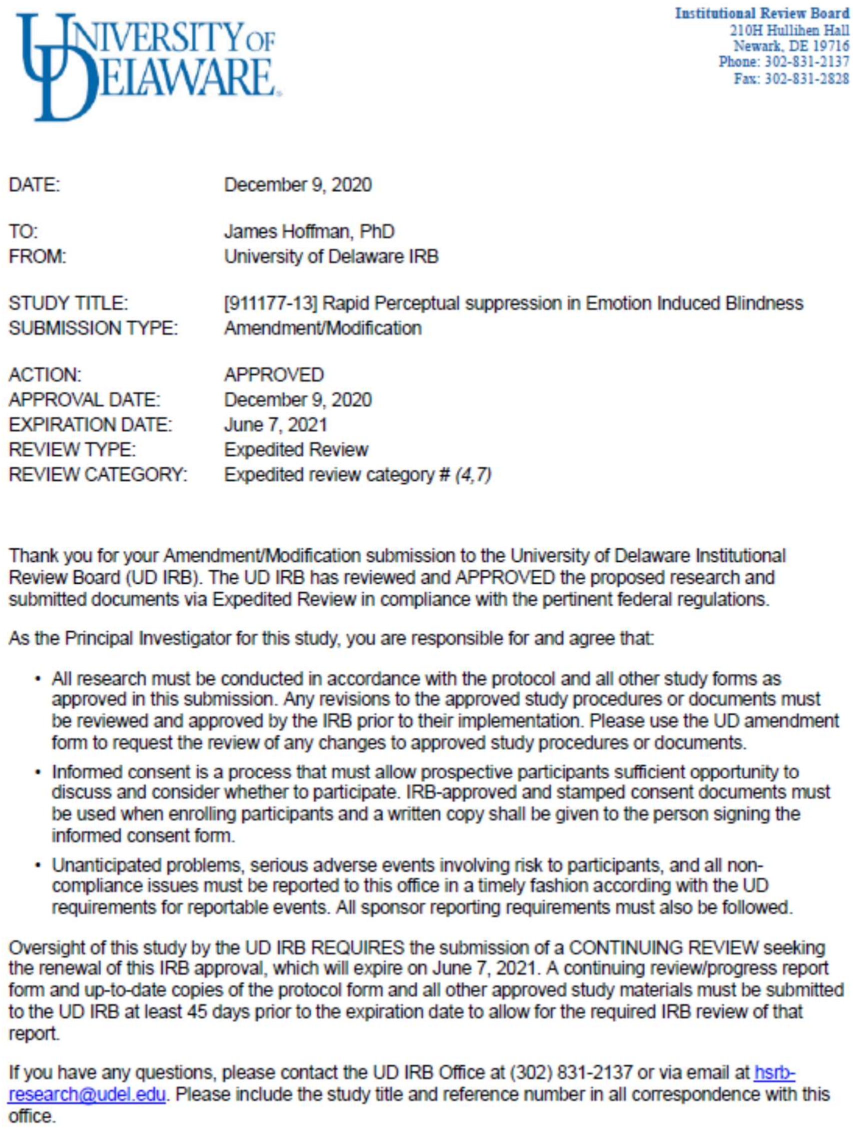


Figure 21. Chart showing the Valence and Arousal of emotional (negative) and neutral distractors. Using a Qualtrics survey and Mturk we gathered data on each large set of emotional and neutral images. We measured emotional valence and arousal of the stimuli on a 9-point Likert scale. Emotional distractors were selected from the large stimulus set if they were above 6 in both valence and arousal. Neutral stimuli were selected if they were between 3 and 5.5 in both valence and arousal. The stimuli significantly differed in valence and arousal ($p < .05$).

Appendix C

IRB/HUMAN SUBJECTS APPROVAL



UNIVERSITY OF DELAWARE

Institutional Review Board
210H Hullihen Hall
Newark, DE 19716
Phone: 302-831-2137
Fax: 302-831-2828

DATE:	December 9, 2020
TO:	James Hoffman, PhD
FROM:	University of Delaware IRB
STUDY TITLE:	[911177-13] Rapid Perceptual suppression in Emotion Induced Blindness
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVED
APPROVAL DATE:	December 9, 2020
EXPIRATION DATE:	June 7, 2021
REVIEW TYPE:	Expedited Review
REVIEW CATEGORY:	Expedited review category # *(4,7)*

Thank you for your Amendment/Modification submission to the University of Delaware Institutional Review Board (UD IRB). The UD IRB has reviewed and APPROVED the proposed research and submitted documents via Expedited Review in compliance with the pertinent federal regulations.

As the Principal Investigator for this study, you are responsible for and agree that:

- All research must be conducted in accordance with the protocol and all other study forms as approved in this submission. Any revisions to the approved study procedures or documents must be reviewed and approved by the IRB prior to their implementation. Please use the UD amendment form to request the review of any changes to approved study procedures or documents.
- Informed consent is a process that must allow prospective participants sufficient opportunity to discuss and consider whether to participate. IRB-approved and stamped consent documents must be used when enrolling participants and a written copy shall be given to the person signing the informed consent form.
- Unanticipated problems, serious adverse events involving risk to participants, and all non-compliance issues must be reported to this office in a timely fashion according with the UD requirements for reportable events. All sponsor reporting requirements must also be followed.

Oversight of this study by the UD IRB REQUIRES the submission of a CONTINUING REVIEW seeking the renewal of this IRB approval, which will expire on June 7, 2021. A continuing review/progress report form and up-to-date copies of the protocol form and all other approved study materials must be submitted to the UD IRB at least 45 days prior to the expiration date to allow for the required IRB review of that report.

If you have any questions, please contact the UD IRB Office at (302) 831-2137 or via email at hsrb-research@udel.edu. Please include the study title and reference number in all correspondence with this office.