Motivationally Significant Stimuli Show Visual Prior Entry: Evidence for Attentional Capture

Greg L. West, Adam A. K. Anderson, and Jay Pratt
University of Toronto

Previous studies that have found attentional capture effects for stimuli of motivational significance do not directly measure initial attentional deployment, leaving it unclear to what extent these items produce attentional capture. Visual prior entry, as measured by temporal order judgments (TOJs), rests on the premise that allocated attention accelerates perception, that is, stimuli that receive attention first are perceived first; thus, this method is a sensitive and direct measure of the time course of initial attentional deployment. The authors demonstrate, using a novel TOJ paradigm without cues, that displays of faces and facial threat show visual prior entry effects, that is, these stimuli are prioritized by the perceptual–attentional system over other stimuli competing for awareness. This study provides direct evidence of the extent to which motivationally significant stimuli capture attention over other concurrently displayed items in the visual array.

Keywords: attentional capture, emotion, temporal order adjustments, biased competition

Of the vast amount of information that reaches the retina, only a small portion is selected for further processing as attentional mechanisms restrict what information is passed on to the higher visual processing centers of the brain. In a very real sense, what we attend to is what we perceive (Serences & Yantis, 2006). Often the orienting of attention across the visual field is goal dependent. For example, when searching for a red pen, our attention system selects red objects, resulting in the tendency to perceive such objects to the exclusion of others. Conversely, certain task-relevant stimuli also have the ability to cause reflexive shifts of attention. For example, while reading a book, the sudden appearance of a fly in the visual field will cause an involuntarily shift of attention regardless of how attention is currently allocated; in other words, the fly will capture attention. Stimuli that capture attention are not only the reflexive recipients of increased amounts of attention, but are also perceived prior to other contemporaneous stimuli. The present studies examined this latter index, examining whether faces and their emotional contents are the very first to receive attention.

Despite top-down influences having been shown to override involuntary shifts of attention in certain situations (e.g., Folk, Remington, & Johnston, 1992; Yantis & Jonides, 1984), several studies have demonstrated that some elementary stimulus features can reflexively capture attention. For instance, rapid changes in luminance have often been reported to involuntarily orient attention (Irwin, Colcombe, Kramer, & Hahn, 2000; Theeuwes, 1995). Similarly, others have demonstrated that abrupt visual onsets also attract attention, even if these events are not related to the task goal (Jonides, 1988; Yantis, 1993; Yantis & Jonides, 1996). It is argued that this is because attention is primarily allocated to the most salient item in the visual array (Theeuwes, 1992, 1995). The attentional system’s ability to quickly detect these irrelevant but salient or novel items in the environment is considered to be adaptive; as such, an item’s potential consequence toward the perceiver is unclear at the time of its appearance.

Salience can also be mediated through more complex environmental stimuli, and such stimuli may also produce reflexive shifts of attention in the same way as more elementary forms of stimuli. For people, an especially good example of this is faces because they provide both socially and biologically relevant information. Facial characteristics are crucial for detecting static properties, such as a person’s age, gender, and identity. Inherently, faces also convey dynamic properties, such as emotion, which is also important given that facial expressions provide critical information about potential threats and rewards present in one’s environment (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Whalen et al., 1998). Thus, it is not surprising that faces, acting as motivationally significant stimuli, are thought to influence the orienting of attention and have been argued to capture attention (e.g., Eastwood, Smilek, & Merikle, 2003; Vuilleumier & Schwartz, 2001b); yet, the degree to which they demand attention over other concurrently displayed stimuli remains unclear.

Using a variety of techniques, researchers have investigated whether faces or emotional faces bias attention; however, no study to date has specifically measured how attention is initially deployed toward these stimuli. Instead, indirect measurements have been used to infer the existence of attentional capture effects. For example, neuropsychological studies, using neglect patients with...
unilateral inattention, suggest that attention is biased toward stimuli of motivational importance. Vuilleumier and Schwartz (2001b), using schematic emotional faces presented to visuospatial neglect patients, reported that faces with a happy or threatening configuration resisted extinction (the failure to perceive objects in the unattended field) more so than shapes, suggesting that these items demand attention over others. Similar findings in neglect patients were found using other threat-related items (e.g., spiders) compared with neutral or positive items (e.g., flowers; Vuilleumier & Schwartz, 2001a). These results are thought to occur because of preserved subcortical limbic pathways with the amygdala acting to modulate visual processing through feedback connections to areas of the visual cortex (Anderson & Phelps, 2001; LeDoux, 2000; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). Although these studies provide evidence that attentional deployment is influenced by subcortical structures related to processing of motivational significance, attentional engagement is not directly measured. Thus, the degree to which these types of stimuli produce attentional capture over others is not yet clearly understood.

A wide variety of experimental methods have also been used in attempts to provide evidence for the ability of both neutral and emotional faces to capture attention in healthy populations. For instance, Ro, Russell, and Lavie (2001) used a change detection paradigm to examine differences in detection advantage for items from different categories. It was found that participants were more accurate in detecting changes between faces than between objects from other categories (e.g., household appliances), allowing the authors to conclude that faces have a special status when competing for selective attention. Similarly, the ability for faces to modulate attention to a larger degree than items from other categories was found by Mack, Pappas, Silverman, and Gay (2002) using a rapid serial visual presentation task. In this study, probe detection was greater when a schematic face was used as a target, compared with when either a tree or an inverted face was used. Negative emotional faces appear to modulate the allocation of attention to an even greater degree. For example, Ohman, Lundqvist, and Esteves (2001) used schematic displays of facial threat in a visual search task, and found that participants were quicker to detect a threatening target among neutral distractors than when the target face had a friendly configuration. Similar findings were reported by Eastwood et al. (2003), who found that the time to count the number of schematic facial features in a search display was longer when the features were embedded in a sad face compared with a positive or neutral face configuration. In addition, Calvo and Esteves (2005), using a modified change detection task, found an increased detection advantage for emotional faces compared with those neutral in valence.

The aforementioned results provide evidence that motivationally relevant stimuli have a significant impact on attentional processes. These studies, however, provide little information about the exact extent to which these stimuli are initially prioritized by the visual–attentional system because initial attentional deployment was not directly measured. Change detection, rapid serial visual presentation, and visual search tasks, for instance, require goal-dependent attentional deployment involved in the searching process, thus conflating reflexive and goal-directed attention processes. In addition, tasks previously used to measure attentional bias toward these stimuli may require several shifts of attention during a given trial because visual search and change detection tasks typically involve the redistribution of attentional focus to multiple spatial locations. In other words, during these tasks, attention is likely reallocated several times before a response is made. Because of this, these paradigms may not provide an adequate measure of the very earliest deployment of attention.

Possible problems with response-time-based visual search paradigms have been noted in a recent review of the visual search literature by Horstmann (2007). He noted that studies (including his own replications) that use visual search to investigate the enhanced processing of negative facial displays (i.e., sad or threatening configurations) rarely find flat search slopes with increasing set size, indicating an absence of reflexive attentional orienting toward these items. Instead, relative decreases in search slopes are found for negative facial targets relative to neutral targets, indicating an increased detection advantage for negative configurations. It may still be possible, however, that negative faces do demand reflexive attentional deployment over other stimuli, but that visual search paradigms are simply not sensitive enough to detect these subtle differences. Typical search items that “pop out” of the visual array, and therefore produce flat search slopes, are typically composed of one basic elementary feature (e.g., size, color, luminance). Therefore, when considering more complex stimuli used in visual search paradigms, the observed increase in search slope may be a result of the increased time needed to integrate the conjunctive features of the face stimulus to garner a proper response (Horstmann, 2007). This may even be the case when the holistic representation of the face stimulus in fact captures attention. Thus, it could be possible that visual search paradigms may not be sensitive enough to determine the exact extent to which a face stimulus, neutral or threatening, demands attention over other competing environmental stimuli.

Aside from visual search, the other commonly used method to investigate how threatening and neutral faces are attended is a variant of the classic spatial cuing task (Posner, 1980). Typically, these tasks use the stimuli of interest (e.g., a schematic angry face configuration) as a cue that predicts a subsequent target’s location. Differences in reaction time (RT) between different face–cue conditions are taken as evidence for their ability to differentially influence shifts of attention. Typical results from these studies show increased measures of RT following invalidly cued probes that are preceded by a threatening face stimulus, as compared with RT measures following invalidly cued probes that are preceded by a neutral face stimulus (e.g., Bradley et al., 1997; Cooper & Langton, 2006; Fox, Russo, Bowles, & Dutton, 2001; Fox, Russo, & Dutton, 2002). Given that cuing studies measure RT or discrimination accuracy (Phelps, Ling, & Carrasco, 2006) in response to the detection of a target and not the cue itself, it is not clear how attention is allocated during the actual display of the face–cue stimulus. Rather, these studies assess a global attentional measure created by gauging a combination of different attentional processes, for example, attentional disengagement, attentional shifting, or attentional reengagement (Posner, Walker, Friedrich, & Rafal, 1984), but do not measure degrees of initial attentional deployment. Fox et al. (2001), in fact, concluded that their differential face–cue findings are due to modulations of attentional disengagement (i.e., an increased holding of attention at the spatial location of an angry face stimulus) and not attentional deployment. Thus, socially and motivationally important events may receive
more or prolonged attention, but it remains unclear whether these events are the very first to attract attention.

Visual prior entry, as assessed by temporal order judgments (TOJs), is a sensitive measure of initial attentional deployment, and thus provides a direct and accurate measure of attentional capture. Prior entry occurs because attention accelerates the processing of sensory stimuli, thereby decreasing the time between the physical onset of a stimulus and its further processing (for review, see Shore, Spence, & Klein, 2001; Stelmach & Herdman, 1991; Ulrich, 1987). In other words, a stimulus cannot be further processed by the visual system without the simultaneous involvement of both attentional and perceptual processes (e.g., Serences & Yantis, 2006); therefore, a stimulus that receives attention first will achieve awareness before other competing stimuli. Typically, TOJ tasks use cues to shift spatial attention to a particular location before the onset of two target items separated by some variable interval. Participants report which target item had the first onset at different stimulus onset asynchrony (SOA) intervals, producing a TOJ response function. A point of perceived stimulus simultaneity (PSS) is then calculated, indicating the interval needed for the participant to perceive both target items as arriving simultaneously. Thus, the PSS shows how much time the unattended stimulus must occur before the attended stimulus, due to prior entry, for them to be perceived as occurring simultaneously. Because of this, the TOJ paradigm, through the PSS, provides a precise measure of the extent of attentional capture.

Past research using TOJ has consistently demonstrated that spatial attention modulates the speed at which information is processed in the visual system. For example, Stelmach and Herdman (1991) demonstrated that participants perceived a black dot at a spatially cued location several milliseconds before the same dot at an uncued location. This PSS of several milliseconds is taken as an indication of the amount of time the uncued dot had to appear before the cued dot so that both items would be subjectively perceived as arriving at the same time. Others have also investigated how low-level visual events affect prior entry. For instance, Shore et al. (2001) found that an exogenous cue (a peripheral flash near the target location) produced a larger prior entry effect for a bar stimulus as compared with an endogenous cue (an arrow pointing to the target location). It is important to note that various control tasks such as ternary response tasks (Ulrich, 1987), simultaneity judgments (Stelmach & Herdman, 1991), and altering task instructions (Shore et al., 2001) have demonstrated that observed prior entry effects are not driven by response bias. In other words, accelerated stimulus perception associated with prior entry has been reliably shown to result from allocated spatial attentional and not some other nonattentional account.

In the present study, to accurately measure initial attentional deployment toward faces both neutral and threatening in configuration, we used a novel TOJ paradigm in which the stimuli of interest were used to directly affect attentional allocation instead of cues. In this task, foveal attention was held at fixation; thus, spatial attention was presumably equally distributed across the initial display, allowing the face stimuli themselves to influence subsequent attentional deployment. Thus, the amount of time that one stimulus receives attention before another could be calculated (i.e., the PSS), indicating the temporal magnitude of attentional capture.

Experiment 1

In Experiment 1, we explored whether a face, a stimulus of social and biological relevance, captures attention by contrasting a face with an inverted face stimulus. Previous work has shown that inverted faces are processed more like nonface objects than faces (e.g., White, 1999; Farah, Wilson, Drain, & Tanaka, 1995) and that their recognition relies more on featural processing, whereas face recognition relies more on holistic processing (e.g., McKelvie, 1995; Bartlett & Searcy, 1993). Consequently, many previous studies have used inverted faces as control stimuli for upright faces because this controls for physical aspects of the face while impairing holistic processing (e.g., Bindermann, Burton, Hooge, Jenkins, & de Haan, 2005; Eastwood, Smilek, & Merkle, 2001; Mack et al., 2002; Phelps et al., 2006; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Ro et al., 2001; Vuilleumier & Schwartz, 2001b). In addition, schematic faces were used instead of real face stimuli because this allowed for the careful control of luminance, contrast, and key physical features used to express different emotions. Previous research has raised concerns over the use of photographic facial stimuli in attentional research because differences in low-level luminance contrast between photographic stimuli have been shown to affect attentional orienting independent of the actual facial configuration (Purcell, Stewart, & Skov, 1996). Indeed, many studies have used schematic faces to avoid this potential confound (Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002; Calvo & Esteves, 2005; Eager, Jedynak, Iwaki, & Skrandies, 2003; Eastwood et al., 2001; Fenske & Eastwood, 2003; Fox et al., 2000, 2001, 2002; Lundyqvist, Esteves, & Öhman, 1999; Mack et al., 2002; Northdurft, 1993; Öhman et al., 2001; Tipples, Atkinson, & Young, 2002; Vuilleumier & Schwartz, 2001b; White, 1994). We reasoned that if faces do capture attention, then an upright face should show a visual prior entry effect when contrasted against an inverted face.

Method

Subjects. Sixteen right-handed undergraduates (13 women) from the University of Toronto participated for course credit. All participants were naïve to the purpose of the study and had normal or corrected-to-normal vision.

Apparatus and procedure. The experiment was conducted on a PC computer with a CRT monitor at refresh rate of 85 Hz in a dimly lit, sound-attenuated room. Viewing distance was held constant at 44 cm with a chin/headrest. Eye movements were monitored using closed-circuit television to ensure task compliance.

A schematic of a typical trial sequence is shown in Figure 1. Each trial began with an initial display consisting of a black (0.0 cd/m²) fixation cross (0.5° × 0.5°) with two placeholder boxes (4.25° × 4.25°), each of which was centered 3.15° from fixation on a white (31.5 cd/m²) background. Participants were instructed to fixate on the cross while allocating attention equally across the display. After 1,000 ms, an onset of either a schematic upright or inverted face, both 3.25° in diameter, occurred in one of the placeholder boxes. After a variable interval (0, 12, 24, 48, 60, or 108 ms), the onset of the remaining schematic appeared in the opposite box. Both schematic faces remained on the screen for an additional 62 ms before being masked in synchrony until response. Participants were instructed to determine which item appeared
first, the face (z key) or the inverted face (/ key). Participants were encouraged to be accurate in their responses while still responding quickly. After a response was made, a blank white background was displayed for 1,000 ms before the next trial began.

**Design.** Each participant completed 20 practice trials with longer intervals between stimulus onsets before completing 396 experimental trials across six blocks. Trial order was randomized across both factors (first onset stimulus and SOA). Both the face and the inverted face were equally likely to have first onset at each nonzero SOA and equally likely to be displayed to the left or right of fixation. Short breaks were given between blocks.

**Results and Discussion**

Trials with RTs greater than 2,000 ms were excluded from analysis (5% of all data). Figure 2a shows the percentage of responses indicating that the face was detected first at a given SOA averaged across participants. Negative SOAs refer to when the face was presented first, whereas positive SOAs indicate when the inverted face was presented first. The effect of prior entry was assessed by calculating each participant’s individual PSS, indicating the interval needed for the participant to perceive both target items as arriving simultaneously. This was derived by determining the intercept at the 50% point on the regression line of each participant’s TOJ function (Frey, 1990; Shore et al., 2001; Stelmach & Herdman, 1991; Ulrich, 1987). The average PSS was 5.88 ms, which was significantly different from zero, $t(15) = 2.31, p < .038$, indicating a prior entry effect for the face.

In the current experiment, attention was, presumably, initially distributed across the two placeholder locations equally. Despite this, the inverted face would need to be displayed almost 6 ms before the upright face for both stimuli to be perceived as arriving at the same time. This prior entry effect indicates that upright faces
capture attention, thereby achieving awareness before other events with similar low-level features.

Experiment 2

The results of the previous experiment provide evidence that faces capture attention, as defined by being the first recipients of attention (i.e., showing a visual prior entry effect). Among faces, it has been previously argued that those with emotional displays have privileged status and thus have greater attention-capturing properties than faces with neutral expressions (e.g., Calvo & Esteves, 2005; Öhman, 2002; Öhman et al., 2001; Vuilleumier & Schwartz, 2001a, 2001b). To examine this with the TOJ paradigm, Experiment 2 used a schematic facial display of threat (i.e., an angry face) and a schematic neutral face. We chose to use an angry face because many other studies investigating the effects of emotion on attention have used these forms of stimuli (e.g., Fox et al. 2000; Lundqvist et al., 1999) and have found robust effects. The facial stimuli were carefully controlled so that the only featural differences between the neutral and angry face were the curvature of the mouth and the angle of the eyebrows (Calvo & Esteves, 2005; Fox et al., 2000, 2001, 2002; Lundqvist et al., 1999; Öhman et al., 2001). All other aspects of the stimuli used were identical.

Method

Subjects. Fourteen right-handed undergraduates (9 women) from the University of Toronto participated for course credit. All participants were naïve to the purpose of the study and did not participate in Experiment 1. Participants had normal or corrected-to-normal vision.

Apparatus, procedure, and design. The apparatus, procedure, and design were the same as in Experiment 1 except that the stimuli used were now schematic angry and neutral faces.
Results and Discussion

Again, trials with RTs greater than 2,000 ms were excluded from analysis (4% of all data). Figure 2b shows the percentage of responses indicating that the angry face was detected first at a given SOA averaged across participants, with negative SOAs referring to when the angry face was presented first and positive SOAs indicating when the neutral face was presented first. Again, the effect of prior entry was assessed by calculating each participant’s individual PSS. The average PSS was found to be 7.85 ms, which significantly differed from zero, $t(13) = 2.59, p < .025$, indicating a prior entry effect for the threatening face.

The current experiment provides strong evidence that a threatening face configuration is attended to and perceived before a contemporaneous neutral face stimulus. As indicated by the observed prior entry effect, a face with a neutral configuration would have to be presented almost 8 ms before a face with a threatening configuration for both items to be subjectively perceived as arriving simultaneously, thus demonstrating that facial displays of threat capture attention.

Experiment 3

It is possible that observed visual prior entry effects for the angry face may not have been driven by its holistic processing but by low-level featural differences between both exemplars (i.e., the curvature of the mouth or the angle of the eyebrows). To avert such concerns, we conducted a third experiment contrasting an angry face with an inverted angry face, thus controlling for distinctive features. Prior studies have shown that facial inversion disrupts the processing of emotional content (e.g., McKelvie, 1995; Phelps et al., 2006; Pourtois et al., 2005). If the effects seen in the previous experiments are due to the emotional content arising from the holistic processing of the angry face, then a prior entry effect similar to that found in the previous experiment should be observed in the current experiment.

Method

Subjects. Fourteen right-handed undergraduates (9 women) from the University of Toronto participated in Experiment 4 for course credit. All participants were naïve to the purpose of the study and did not participate in any previous prior entry experiments. Participants had normal or corrected-to-normal vision.

Apparatus, procedure, and design. The apparatus, procedure, and design were the same as in previous experiments except that the stimuli used were now a schematic inverted angry face and a schematic inverted neutral face.

Results and Discussion

As in the previous experiments, trials with RTs greater than 2,000 ms were excluded from analysis (5% of all data). Figure 2c, with negative SOAs referring to when the angry face was presented first and positive SOAs indicating when the inverted angry face was presented first. The average PSS was 7.17 ms, which differed significantly from zero, $t(15) = 2.46, p < .025$, once again indicating a prior entry effect for the threatening face.

The finding from the present experiment corroborate with the previous, as the threatening face configuration produced a prior entry effect similar in magnitude as that found in Experiment 2. As this experiment controlled for featural aspects of the angry face, we can attribute the prior entry effects observed in the previous experiments as being the result of the holistic processing of the face and its emotional content and not a result of low-level featural stimulus differences.

Experiment 4

To further investigate whether the prior entry effect found in Experiment 2 was due to low-level featural differences between the neutral and angry face, we conducted a fourth experiment that inverted both face stimuli. Again, given that inverting a face stimulus breaks down its holistic processing and the efficient extraction of its emotional contents, prior entry effects should be eliminated in the present experiment if results from Experiment 2 were in fact driven by holistic emotional processing. If, however, previous results were due to low-level featural confounds, then a prior entry effect should still be observed. In addition, if participants still report that the inverted angry face produces prior entry, this could also indicate a response bias present in previous experiments where participants simply indicate the more “interesting” item as having first onset when uncertain. The elimination of a prior entry effect under the current experimental conditions would also discount this possibility.

Method

Subjects. Fourteen right-handed undergraduates (9 women) from the University of Toronto participated in Experiment 4 for course credit. All participants were naïve to the purpose of the study and did not participate in any previous prior entry experiments. Participants had normal or corrected-to-normal vision.

Apparatus, procedure, and design. The apparatus, procedure, and design were the same as previous experiments except that the stimuli used were now a schematic inverted angry face and a schematic inverted neutral face.

Results and Discussion

As in the previous experiments, trials with RTs greater than 2,000 ms were excluded from analysis (4% of all data). Figure 2d shows the percentage of responses at a given SOA averaged across participants, with positive SOAs referring to when the inverted angry face was presented first and negative SOAs indicating when the inverted neutral face was presented first. The average PSS was calculated to be 1.28 ms, which was nonsignificant, $t(13) < 1$, indicating no reliable prior entry effect for either of the inverted face stimuli. The elimination of the prior entry effect in the current experiment further confirms that the results observed in Experiment 2 were in fact driven by the holistic processing of both facial stimuli, resulting in their differential effect on attentional engagement. Thus, these results further indicate that prior entry for the angry face can be attributed to the processing of emotional content and not low-level featural salience. In addition, the current experiment suggests that previous results were not driven by a response bias toward reporting the angry face as arriving first under conditions of uncertainty. The current results also address the concern that one of the two items displayed will always have a prior entry effect. Instead, it appears that attentional
capture as indexed by prior entry is reserved for stimuli of social and biological significance.

Experiment 5

The previous experiments in the current study used schematic face stimuli because they allow for the control of low-level featural aspects and luminance levels that might otherwise produce their own capture effects. This way we are reasonably certain the previous visual prior entry effects are due to the specific facial features only. It is possible, however, that the previous prior entry effects for neutral and threatening face configurations were artifacts resulting from the use of schematic stimuli. Indeed, some research has suggested that it is possible that schematic face stimuli produce greater behavioral effects as compared with photographic face stimuli (Hietanen & Leppänen, 2003). Furthermore, the use of schematic stimuli may limit the ecological validity of our present findings. We therefore wanted to investigate whether our previously observed effects could be extended to real human faces. Accordingly, Experiment 5 contrasted human facial displays of anger (i.e., threat) with faces neutral in valence. Because the small changes in the features of the schematic faces reliably produced differential attentional capture effects, we predicted that similar results would be found when using human facial displays of threat. Thus, a realistic threatening face should show visual prior entry when contrasted against a realistic neutral face. Furthermore, to minimize the effects of low-level featural differences, we matched the luminance and contrast of threatening and neutral displays.

Method

Subjects. Twelve right-handed undergraduates (7 women) from the University of Toronto participated in Experiment 5 for course credit. As in previous experiments, all participants were naïve to the purpose of the study and did not participate in any previous prior entry experiments. Participants had normal or corrected-to-normal vision.

Apparatus, procedure, and design. The apparatus, procedure, and design were the same as in previous experiments except that the stimuli used were now black and white photographs of facial displays of anger and faces neutral in valence. The photographs were enclosed in a 3.80° × 3.60° border excluding most of the hair and nonfacial contours (see Figure 1). Because of variability in real facial stimuli, a larger set of face photographs was rated on factors of induced arousal and perceived emotional valence. Four angry face and four neutral face stimuli were chosen (two female and two male for each category) on the basis of interrater reliability of valence and arousal levels of stimuli within each category. Contrast and luminance levels were adjusted across all photographs to ensure as much similarity as possible across these factors (Vuilleumier, Armony, Driver, & Dolan, 2003). Every trial contrasted a threatening face with a neutral face, with each face exemplar from one category being equally likely to be paired with each face exemplar of the opposite category. All other aspects of apparatus and design were identical to previous experiments.

Results and Discussion

As before, trials with RT greater than 2,000 ms were excluded from analysis (4% of all data). The results are shown in Figure 3a, with negative SOAs referring to when the angry face was presented first and positive SOAs indicating when the neutral face was presented first. The average PSS was 18.26 ms, which differed significantly from zero, t(11) = 3.78, p < .003, showing that prior entry for facial displays of threat extend to human face stimuli. To examine the effect of real versus schematic faces, when comparing the results of Experiment 2 (schematic angry face vs. schematic neutral face) with the present experiment, it appears that realistic human facial displays of threat produce a significantly greater degree of prior entry when compared to schematic displays of threat, t(11) = 2.37, p < .037 (see Figure 4). Thus, the prior entry for angry faces was not an artifact of using schematic stimuli as the current experiment indicates that photographic displays of facial threat appear to produce consistent or even greater degrees of visual prior entry.

Experiment 6

There still remains the possibility that previously observed prior entry effects may be due to a response bias toward reporting the more interesting or threatening item as having first onset under conditions of uncertainty. The results of Experiment 4 already begin to discount this possibility, but a conclusion based on that evidence might be confounded by referring to the angry face as an “inverted angry face” during task instructions, or it is also perhaps a possibility that such a response bias may be present only for upright faces. We therefore conducted a sixth experiment to investigate this possibility, using the stimuli from Experiment 5 because these items produced the largest PSS, but with changing the task instructions so that any present response bias of this nature would have an opposite effect on the direction of reported prior entry. Specifically, participants were asked to report which face stimulus had the second onset. Thus, if previous effects were primarily driven by a response bias toward simply reporting that the threatening face had the first onset when uncertain, an opposite response pattern should be observed under these new task conditions (Shore et al., 2001) because participants would now report the angry face as having second onset when uncertain. If, however, prior entry for motivationally significant stimuli is in fact attentionally driven, a consistent PSS in the same direction as previous experiments will be observed.

Method

Subjects. Fourteen right-handed undergraduates (8 women) from the University of Toronto participated in Experiment 6 for course credit. Again, all participants were naïve to the purpose of the study and did not participate in any previous prior entry experiments. Participants had normal or corrected-to-normal vision.

Apparatus, procedure, and design. The apparatus, procedure, design, and stimuli were the same as in Experiment 5 with the exception of the instructions given to the participant. Participants were now instructed to determine which facial stimulus appeared second.

Results and Discussion

As in previous experiments, trials with RTs greater than 2,000 ms were excluded from analysis (6% of all data). Figure 3b shows the percentage of responses indicating that the angry face was
perceived first at a given SOA. Again, negative SOAs refer to when the angry face was presented first, whereas positive SOAs indicate when the neutral face was presented first. The average PSS was calculated to be 7.66 ms, which differed significantly from zero, \( t(13) = 2.85, p < .017 \), again demonstrating prior entry for human angry faces. If previous prior entry effects were primarily due to a response bias toward selecting the angry face under conditions of uncertainty, then we should have observed a PSS in the opposite direction for the “which second” task. Instead, we found that the PSS remained in the same direction as previous experiments, indicating that the angry face was still attended first. The reduction in PSS magnitude in Experiment 6 compared with Experiment 5 under the new task condition was nonsignificant, \( t(24) = 1.13, p < .20 \) (see Figure 4). It is nonetheless probable that some amount of response bias was present in previous experiments. This relative reduction in PSS magnitude between “which first?” and “which second?” tasks was also observed by Shore et al. (2001), indicating that classic TOJ tasks that use exogenous cues to shift attention also have some degree of response bias present within the reported PSS. The present results nonetheless confirm that prior entry for complex stimuli exists independent of response bias, as the direction of the PSS still indicated that the angry face was attended first even when a counteracting response bias was introduced. Thus, we can be confident that previously observed prior entry effects are in fact primarily attentionally driven and not due to some other nonattentional account.

**General Discussion**

The current results clearly demonstrate that stimuli of motivational significance are the first recipients of attention and thus are shown to capture attention. Visual prior entry effects, which directly reflect attentional capture, were found when using both schematic and realistic human faces as stimuli. Specifically, Experiment 1 demonstrated that a neutral face produces prior entry when contrasted against an inverted face, thus receiving initial attentional deployment and capturing attention. When a facial display of threat is contrasted with a neutral face, however, as in Experiment 2, the threatening face configuration then produces visual prior entry. Of noteworthiness, this effect has also recently been shown by Fecica and Stolz (2008), using schematic facial displays. It is important to note that this effect was found to be due to the processing of their threat-based emotional content and not low-level featural differences (Experiments 3 and 4), and was also found to extend to the use of photographic face stimuli (Experiment 5). Also important, Experiments 4 and 6 provided strong evidence against the possibility of a response bias being primarily

![Figure 3](image-url)  
**Figure 3.** The average proportions of angry face first responses (“which first?” task; A) and neutral face second responses (“which second?” task; B) are displayed as a function of stimulus onset asynchrony (SOA) between both contrasted stimuli in each experiment. Photographic angry faces were contrasted with photographic neutral faces. Negative SOAs indicate that the angry face was presented first, whereas positive SOAs indicate that the neutral face was presented first. The solid vertical line shows the zero SOA trials (simultaneous onset of both stimuli), with the horizontal line showing the 50% response mark. The dashed vertical line to the right of the arrows represents the actual perceived stimulus simultaneity calculated from the temporal order judgment function. The visual prior entry effect is shown as the displacement of the dashed line from the solid vertical line.

![Figure 4](image-url)  
**Figure 4.** Visual prior entries in milliseconds are displayed for Experiments 1 through 6, representing the observed attentional capture effects. Error bars represent 1 standard error. Inv. = inverted.
responsible for driving these effects. Previous research has demonstrated that facial stimuli (Bindermann et al., 2005; Ro et al., 2001) and emotional stimuli (e.g., Anderson, 2005; Fox et al., 2001, 2002; Lang, Bradley, & Cuthbert, 1997; Most, Chun, Widders, & Zald, 2005; Öhman et al., 2001) can affect various attentional processes. The present study directly measured the initial deployment of attention to these stimuli, thereby demonstrating how stimulus significance may bias the contents of awareness by being the very first recipients of attention. In addition, the current results provide a direct measure of the temporal offset of these capture effects.

The results from Experiment 2, using the TOJ paradigm, reveal that when competition for attentional resources exists between a threatening stimulus and one that is also behaviorally important but nonthreatening, the threat-based stimulus is perceived to occur prior in time. This prior entry occurs even when a neutral stimulus is actually presented first, providing evidence that facial displays of threat capture attention over neutral displays, with this prior attentional engagement accelerating their processing, thereby decreasing the time between the physical onset and perception. An emotional item’s ability to produce prior entry is complementary with recent findings, suggesting that facial expressions enhance subsequent spatial attention (Phelps et al., 2006; Pourtois et al., 2005) as these studies exhibit greater or prolonged, but not faster, attentional allocation.

The present findings also fit well into the existing neurological literature regarding face and emotion processing. The existence of reciprocal subcortical connections between the amygdala, parietal regions, and primary and secondary visual cortex (LeDoux, 2000; Morris, Öhman, & Dolan, 1999) allows for the expedient detection of coarse affective properties of stimulus items (Anderson et al., 2003; Anderson & Phelps, 2001). Direct projections to the amygdala ranging from the thalamus to higher-order anterior ventral visual cortices provide ample anatomical evidence that the brain is prepared to reflexively detect emotionally significant items in the environment at the expense of others. Indeed, neuroimaging and lesion studies have implied that the amygdala undertakes a rapid and coarse analysis of emotionally arousing stimuli (Anderson et al., 2003; Głäscher & Adolphs, 2003; Vuilleumier et al., 2003; Whalen et al., 1998), suggesting that subsequent attentional deployment is reflexively allocated to these items (Anderson & Phelps, 2001). In addition, this system has been shown to be active even when emotional faces are not perceived because of backward masking (Whalen et al., 1998) or binocular suppression (Williams, Morris, McGlone, Abbott, & Mattingley, 2003), thus providing a potential mechanism for the attentional system’s prioritization of emotional content observed in the current study.

Consistent with findings of context-dependent capture effects using elementary stimulus features (e.g., Folk et al., 1992; Yantis & Jonides, 1984), the present study suggests that face and threat-based capture effects are also context dependent. Thus, a given stimulus may or may not maintain the ability to capture attention depending on what degree of importance it possesses relative to competing stimuli. In the case of the current study, the most salient item in the visual array (i.e., facial threat) demands initial attentional engagement over other items, thus capturing attention (Theeuwes, 1992, 1995). This is consistent with Desimone and Duncan’s (1995) biased competition theory of attention: Both top-down attention and, in the case of the present study, bottom-up stimulus significance resolve competition between stimuli competing for awareness.

The results of the current study also speak to the use of schematic stimuli in attentional experiments. On the one hand, prior entry effects were found with schematic faces even though the featural differences were relatively subtle. For example, the physical differences between the angry and neutral faces in Experiment 2 were remarkably small (i.e., small changes in curvature of mouth and angle of eyebrows), yet attentional capture for the angry face configuration was still produced. Combined with the findings from other studies (e.g., Lundqvist et al., 1999; Öhman et al., 2001), it appears that schematic faces are entirely suitable as stimuli to investigate issues of attentional allocation. On the other hand, the results of Experiment 5 suggest that there may be a difference in the magnitude of prior entry produced by schematic and photographic faces, namely, prior entry was greater for the photographic displays of facial threat (Experiment 5) when compared with its schematic counterpart (Experiment 2). Thus, depending on the sensitivity of the paradigm being used, realistic faces may be more appropriate. In addition, the finding of greater capture for photographed faces is somewhat at odds with research investigating other forms of attentional orienting produced by facial stimuli. For example, Hietanen, and Leppänen (2003) demonstrated that a spatial cuing effect by gaze direction was stronger when schematic faces were used compared with photographic face stimuli. Thus, it is possible that schematic and real face stimuli differentially affect separate attentional processes, such as attentional disengagement, attentional shifting, or attentional reengagement (Posner et al., 1984). In addition, diminished prior entry for schematic threatening faces may merely reflect that they may less well match real threatening face prototypes.

With regards to the study of prior entry, the TOJ paradigm, until now, has been used to investigate the effect of low-level stimulus features (e.g., abrupt onsets, changed in luminance or contrast) on directed attention (through cues or eye movements) and the speed of perceptual processing. Through these efforts, directed attention has been confirmed to increase the speed of perception (Frey, 1990; Stelmac & Herdman, 1991; Ulrich, 1987), although the rate of perception is dependent on the attentional orienting method used (e.g., exogenous vs. endogenous; Shore et al., 2001). The present study extends this paradigm to the use of “nondirected” attention and more complex stimuli. By doing so, we have expanded the paradigm from answering questions about “what happens when attention is here?” to “where does attention go?”

In the present study, we did not test whether our effects extend to other categories of emotional content (e.g., sad, happy, etc.); therefore, we cannot confirm whether one emotional category would capture attention over another. Previous research suggests that there may be no differential capture effects when faces of different emotions are pitted against each other (e.g., Calvo & Estévez, 2005; Williams et al., 2003). Interestingly, recent findings from Fecica and Stolz (2008) using a similar TOJ paradigm suggest that positive facial affect might be prioritized over negative facial affect. This question continues to be investigated through ongoing research. There also remains the concern that the current set of results might reflect accelerated stimulus perception that is independent of attention. In other words, displays of facial threat might simply be detected faster than other concurrently displayed stimuli. Indeed, there is evidence that emotional faces
can be detected at lower perceptual thresholds compared with neutral faces (Calvo & Esteves, 2005). This, however, does not necessarily mean that attention is absent from this process. According to Serences and Yantis (2006), perception and attention are part of a simultaneous and coordinated group of processes involving multiple visual areas that result in awareness. Whether a stimulus achieves awareness is determined by simultaneous processes that encode low-level visual features, incorporate top-down goal states, and select stimulus categorical representation. In other words, perception and attention can be considered to be part of the same function that selects stimuli for further processing. Both our present results and those of Calvo and Esteves (2005) are consistent with this theoretical perspective; the demonstration of an emotional face’s ability to be detected at a lower perceptual threshold and to produce visual prior entry are probably due to the same underlying process of perceptual acceleration that simultaneously engages attention, thus resulting in the prioritization of the stimulus over other competing items in the visual array.

One other aspect of the current study that should be discussed involves the separation of top-down and bottom-up attention. Although there exists a large body of evidence that suggests that the accelerated perception of motivationally significant stimuli would be driven by bottom-up stimulus salience maps (e.g., Anderson et al., 2003; Glaßcher & Adolphs, 2003; Vuilleumier, Armony, Driver, & Dolan, 2001; Whalen et al., 1998), it is still possible that these observed effects are due to influences from learned top-down cognitive salience. Relatedly, in experimental situations in which an upright face was contrasted with an inverted face, it is possible that participants were deliberately searching for the face in display, thus leaving a top-down attentional account for the prior entry effects observed under these conditions. Further research on the time course of stimulus processing using electrophysiological correlates of onset detection is currently being conducted to investigate this possibility.

In sum, the current study used no cues and allowed the stimuli themselves to covertly orient attention from central fixation. The TOJs thus served as a sensitive measure of initial attentional deployment that was independent of other attentional processes (e.g., attentional disengagement, attentional reengagement). The TOJ task, therefore, has proven to be a sensitive and elegant paradigm for use in attentional research, and may prove useful in future efforts directed at dissociating capture effects from different attentional processes produced by other forms of stimuli. By examining prior entry through TOJs, we have demonstrated that attentional capture is not only reserved for elementary stimuli, but also extends to more complex stimulus events that are biologically and socially significant.

References


Received July 12, 2007
Revision received June 6, 2008
Accepted June 10, 2008