

Faces capture attention: Evidence from inhibition of return

Jan Theeuwes and Stefan Van der Stigchel

Vrije Universiteit Amsterdam, The Netherlands

The human face is a visual pattern of great social and biological importance. While previous studies have shown that attention may be preferentially directed and engaged longer by faces, the current study presents a new methodology to test the notion that faces can capture attention. The present study uses the occurrence of inhibition of return (IOR) as a diagnostic tool to determine the allocation of attention in visual space. Because previous research suggested that IOR at a location in space only occurs after attention has been reflexively moved to that location, the current finding of IOR at the location of the face provides converging support for the claim that faces do have the ability to summon attention.

The human face constitutes one of the most important stimuli for social interactions. In addition, face perception is considered to be the most developed visual perceptual skill in humans. Research using single cell recording (Perrett, Hietanen, Oram, & Benson, 1992) neuroimaging (Kanwisher, McDermot & Chun, 1997), and neuropsychology (Warrington & James, 1967) indicates the importance of face perception by demonstrating that there are specialized brain areas that selectively respond to faces. Behavioural evidence shows that attention may be preferentially directed to faces rather than to other objects in a scene (Ro, Russell, & Lavie, 2001). Even though it may feel intuitive that our attention is pulled to the only face in a scene, there is little evidence that attention is indeed automatically and exogenously captured by a face.

Please address all correspondence to: Jan Theeuwes, Cognitive Psychology, Vrije Universiteit, Boechorststraat 1, 1081BT Amsterdam, The Netherlands. E-mail: J.Theeuwes@psy.vu.nl

We thank Tony Ro for providing the photographs. We thank Mieke Donk, Donald Hoffman, Chris Olivers, and Philippe Schyns for helpful comments on the manuscript. This research was funded by a grant from NWO (Netherlands Organization for Scientific Research), no. 402-01-630-PROG to JT.

The typical approach to determine whether faces capture attention in an exogenous way is to use modifications of the classic spatial cueing paradigms developed by Posner and colleagues (see e.g., Posner, 1980; Posner & Cohen, 1984). For example, in target detection tasks, location cues can facilitate response times (RTs) depending on whether or not the cue is a valid indicator of the impending target's location (e.g., Posner, Snyder, & Davidson, 1980). Whereas endogenous orienting (which is considered to be under volitional control) is studied with predictive cues (e.g., a central arrow predicts with 80% validity the location of the upcoming target), exogenous orienting is studied with noninformative peripheral cues (e.g., a sudden onset stimulus that does not predict which location will contain the target). When the target happens to be presented at the location of the exogenous cue, RTs to the target stimulus are relatively fast; when the cue is invalid RTs are relatively slow. It is widely assumed that these RT costs and benefits reflect the operation of a reflexive orienting system that cannot be affected by top-down set (e.g., expectation and biases).

This very same idea is used in studies investigating whether faces can capture attention. For example, studies that used (emotional) faces as nonpredictive cues have demonstrated that attention is shifted to the location that contains a threatening face (e.g., Fox, Russo, Bowles, & Dutton, 2001; Mathews, Mackintosh, & Fulcher, 1997). Typically, in these studies a neutral, happy, or angry face is presented as a cue to the left or the right of a fixation point. Immediately following the presentation of the cue, a target (a probe dot) is presented at the location of the face or at the location that did not contain an object. Typically, these results show a cue validity effect: RTs to targets presented at the location of the face are faster than those presented at the location that did not contain an object. Moreover, these effects are stronger for angry relative to happy faces (e.g., Fox et al., 2001). One conclusion of this type of experiments is that attention is captured by the face similar to the way attention is captured by a true exogenous event such as an abrupt onset. Because the face pulls attention to its location, the processing of subsequent target at that location is facilitated. Even though some (e.g., Mathews et al., 1997; Vuilleumier, 2002; Yiend & Mathews, 2001) have suggested that this paradigm reveals true attentional capture, it should be realized that the cueing effect observed in these studies may be the result of the inability to disengage attention from faces once attention is focused at the location rather than the result of attentional capture (see, e.g., Fox et al., 2001 for this explanation). In other words, these cueing experiments may not speak to the issue of attentional capture at all. Even though it is widely believed that (emotional) faces capture attention (e.g., Bradley et al., 1997; Cauquil, Edmonds, & Taylor, 2000; Mathews et al., 1997; Roskos-Ewoldsen & Fazio, 1992; Vuilleumier, 2002), there may be in

fact be not much, if any, convincing evidence that faces capture attention in an truly exogenous way.

In a recent study, Ro et al. (2001) used a change blindness paradigm and showed that changes in faces were detected more rapidly and more accurately than changes in other objects in a scene. Ro et al. concluded that in competition with other objects, participants prefer to look at faces rather than at other objects. Because of this preferential attention to faces, a change in a face was more readily detected than a change in another object. Again, as Ro et al. (2001) point out, their study only says something about the extent to which participants prefer to direct attention to faces and does not say anything about the extent to which faces are able to capture attention exogenously.

Given this ambiguity in results, assumptions, and claims, we aimed at finding a new methodology that can be used as a diagnostic to test the notion that faces capture attention. In the current study, we used the phenomenon of “inhibition of return” (IOR; Posner & Cohen, 1984). The basic finding of IOR is that after attention is shifted to a location in space, there is delayed responding to stimuli subsequently displayed at that location (Klein, 2000). One characteristic of IOR is particularly important. The occurrence of IOR at a location in space only follows after attention has shifted *reflexively* to that location. Posner and Cohen observed IOR following the reflexive allocation of attention to a location that contained a brief luminance onset flash. In contrast, when spatial attention was allocated voluntarily to a location in space, IOR was no longer observed. In other words, if IOR occurs it is thought to be the result of an involuntary (or obligatory) shift of spatial attention (Posner & Cohen, 1984; see also Pratt, Sekuler, & McAuliffe, 2001).

Similar to earlier studies (e.g., Fox et al., 2001; Mathews et al., 1997), the present study also used a cueing paradigm. Rather than presenting only one object as a cue (which generates exogenous capture simply because of the asymmetric onset), we presented two objects (either a face or a nonface), one on either side of fixation. The objects were flashed for a brief time. After a variable SOA, participants had to make a speeded saccade to either the location that previously contained either a face or a nonface. If the face captures attention in an exogenous way, we expect IOR at the location that previously contained the face. If there is only an endogenous preference to select a face over a nonface, we do not expect IOR. Note that an additional advantage of the current paradigm is that participants do not receive a particular instruction with respect to the objects. There is no attentional set other than the instruction to make a speeded saccade in the direction of the pointer.

METHOD

Eight observers (two males) completed 96 trials (10 practice trials) on which they had to make a speeded eye movement (see Figure 1). Observers were seated 75 cm from a computer screen with their head positioned on a chinrest. Eye movements were measured by means of an Eye Link eye tracking system with 500 Hz sampling rate. Participants were required to fixate the centre of the screen. After 1000 ms, on either side of fixation at 6.9° eccentricity a black- and-white outline photograph ($2.1^\circ \times 2.1^\circ$ visual angle) was presented for 200 ms. After a variable interval between 600 and 800 ms a central pointer was presented. Upon presentation of the pointer participants made a speeded eye movement to the location indicated by the central arrow. We measured saccade latency. We used photographs from Ro et al. (2001), consisting of six different female faces and six different appliances. In each display there was a photo of a face and a photo of an everyday appliance (fan, mixer, iron, stove, toaster, and phone). Observers were instructed to make a fast saccade after the presentation of the pointer. Nothing was said about the purpose of the photographs.

RESULTS

A paired t -test showed that saccade latencies towards the location that previously contained the face were significantly longer (293 ms) than the location that previously contained another object (282 ms), $t(7) = 2.90$; $p < .05$. This indicates a small but robust IOR effect of 11 ms. A subsequent control experiment involving eight new participants with inverted faces and objects showed no saccade latency differences (inverted face [261 ms] versus inverted other object [260 ms]), $t(7) = 0.52$; $p = .618$ suggesting that the IOR effect for faces is not due to low-level feature differences but is due to the semantic processing of faces. A between-experiment analysis showed no main effects. This analysis only showed a significant interaction between experiment (upright/ inverted photographs) and type of object (face/nonface), $F(1,14) = 3.70$; $p = .037$ (one-sided), confirming that the IOR effect only occurred with upright photographs.

DISCUSSION

The present study used the occurrence of IOR as a diagnostic tool to determine whether faces capture attention. Previous research has suggested that in order for IOR to occur, attention has to first *reflexively* shift to a particular location and subsequently disengage from that location. Only in these circumstances, there is delayed responding for a target presented at the

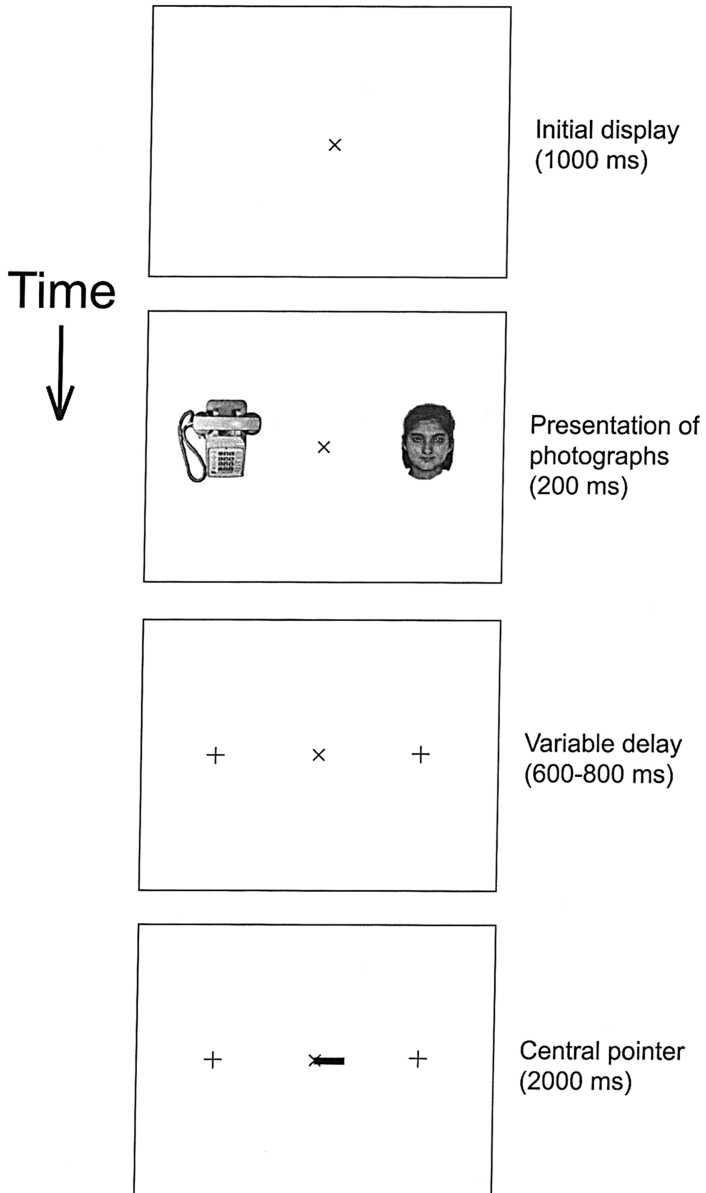


Figure 1. Typical task sequence. Observers fixated the centre fixation point. Two photographs (a face and an appliance) were presented at 6.9° from the fixation point at either side for 200 ms. After a variable interval of 600–800 ms a centre pointer indicated the location to which an eye movement had to be made. Upon the presentation of the centre pointer observers made a speeded saccade to either the location that previously contained a photograph of a face or a photograph of an appliance. Saccade latencies were measured.

location from which attention was previously disengaged (e.g., Pratt et al., 2001; Theeuwes & Godijn, 2002). Since we observed delayed responding of a saccade to a location that previously contained a face, the current findings provide converging support for the notion that faces can summon attention.

The paradigm employed in the current study has many benefits over paradigms used in various studies which claimed to have found evidence for exogenous capture of attention by faces, emotional faces, treating objects, and words. First, in several previous studies only a single object was presented as a cue on only one side of fixation. For example, Fox et al. (2001) presented in each display a single object (e.g., a normal or a jumbled face) as a cue either 8° to the left or the right of fixation. This cue was followed by a probe that required a speeded response. Obviously, the appearance of the single object in the periphery which constitutes an asymmetric luminance onset will capture attention in an exogenous way (e.g., Theeuwes, 1991) regardless of what the object is (whether it is a face or a jumbled face). In the current study, there was no asymmetric luminance change because two objects, one on each side of fixation, were presented simultaneously. Second, in many studies the exogenous capture of attention is inferred from reaction time benefits in responding to an irrelevant probe dot presented at the location of the (emotional) face relative to a non-face. Even though this result is often explained to represent attentional capture (e.g., Mathews et al., 1997) it only shows that once attention has moved to the location of the (emotional) face it is harder to disengage attention from that location relative to a location that contains a nonface (see Fox et al., 2001 for a similar explanation). Since we used the occurrence of IOR as a marker for exogenous capture of attention this concern does not apply to the current study. Third, if two objects (for example a face and a nonface) are presented at either side of fixation and attention resides at one of the locations, this does not necessarily imply that attention was captured. Indeed, it is very well likely that attention goes to both locations sequentially and participants *choose* to focus their attention at the location that contains the face. This does not mean that attention was captured by the face; it only shows that participants prefer looking at faces rather than at other nonface-like objects. Fourth, in almost all studies that involve cueing with (emotional) faces, there is no adequate control for the occurrence of eye movements. In all studies, the instruction to the participants is to not move their eyes. Even though participants may be able to follow these instructions in most of the trials, it is likely that in a subset of trials, the eyes move to the location of the face. Obviously, moving the eyes to one location gives large RT costs and benefits for detecting probes, which may be the result of differences in retinal acuity rather than differences in the distribution of attention. In the current study we measured eye movements and those trials in which participants moved

their eyes to one of the pictures were removed from the analysis. Fifth, all cueing studies that have used RT to a probe to determine the spatial distribution of attention had to use some artificial additional task such as the discrimination of a probe stimulus that could appear at any of the locations in the periphery. It is feasible that the additional task in which participants had to make a subsequent discrimination to a stimulus presented in the periphery may have altered their allocation of attention. In the current study, we did not need to use an additional discrimination task to infer the distribution of attention. The only task for our participants was to make an eye movement to the left or the right of central fixation indicated by a central marker.

Even though it is clear that there are too many substantial concerns with the classic probe dot methodology to infer attentional capture, the technique has been used in a wide variety of studies investigating capture by emotional and threatening words, faces, and objects in a whole range of clinical and nonclinical groups (e.g., Amir, Elias, Klumpp, & Przeworski, 2003; Georgiou et al., 2005; Koster, Verschuere, Crombez, & van Damme, 2004; Schützwohl & Borgstedt, 2005). The current methodology may present an improvement over previously used paradigms. More specifically, the current paradigm can distinguish between effects that are the result of attentional capture and those related to difficulty in disengaging attention. It is feasible that the current paradigm may help in resolving controversies with respect to whether negative, emotional, or threatening faces, objects, and words are able to capture attention in a truly exogenous way (see Fox et al., 2001).

Our findings are in line with earlier studies that show that faces have to ability to grab attention. In order to detect faces outside the direct focus of attention, it is required that there are perceptual processes that automatically scan and analyse the visual field for face stimuli. Because faces capture attention one has to assume that faces are discriminated by some "preattentive": or unconscious processing. Upon discerning the face, focal attention is directed towards the location of the face in an automatic way. Because a face is an evolutionary, social relevant stimulus it may receive priority for processing, similar to other types of critical events such as luminance onsets and the sudden appearance of a new object (Posner, 1980; Theeuwes, Kramer, Hahn, & Irwin, 1998). One, rather speculative way, to reconcile these findings is to assume a fast pathway for face processing that can circumvent the cortex. It is known that, among other structures such as the Fusiform Face Area (FFA; e.g., Kanwisher et al., 1997), the amygdala plays an important role in face processing (Haxby, Hoffman, & Gobbini, 2000). It is suggested that there may be two parallel pathways to the amygdala: One subcortically and one cortically mediated pathway (LeDoux, 2000). In line with this notion, our finding suggests that by circumventing

the cortex, the “preattentive” thalamo-amygdala pathway can process information fast and in an automatic manner. Upon discriminating the face, attention is automatically shifted to the location of the face, allowing the engagement of cortically mediated pathways (including FFA) for detailed processing of the stimulus.

REFERENCES

- Amir, N., Elias, J., Klumpp, H., & Przeworski, A. (2003). Attentional bias to threat in social phobia: Facilitated processing of threat or difficulty disengaging attention from threat? *Behaviour Research and Therapy*, *41*, 1325–1335.
- Bradley, B. P., Mogg, K., Millar, N., Bonham-Carter, C., Fergusson, E., Jenkins, J., & Parr, M. (1997). Attentional biases for emotional faces. *Cognition and Emotion*, *11*, 25–2.
- Cauquil, A. S., Edmonds, G. E., & Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *NeuroReport*, *11*, 2167–2171.
- Fox, E., Russo, R., Bowles, R. J., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in sub-clinical anxiety? *Journal of Experimental Psychology: General*, *130*, 681–700.
- Georgiou, G.A., Bleakley, C., Hayward, J., Russo, R., Dutton, K., Eltiti, S., & Fox, E. (2005). Focusing on fear: Attentional disengagement from emotional faces. *Visual Cognition*, *12*, 145–158.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, *4*, 223–233.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.
- Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Science*, *4*, 138–147.
- Koster, E. H. W., Crombez, G., Verschuere, B., & de Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*, 1183–1192.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.
- Mathews, A., Mackintosh, B., & Fulcher, E. P. (1997). Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences*, *1*, 340–345.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London: Series B. Biological Sciences*, *335*, 23–30.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160–174.
- Pratt, J., Sekuler, A. B., & McAuliffe, J. (2001). The role of attentional set on attentional cueing and inhibition of return. *Visual Cognition*, *8*, 33–6.
- Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science*, *12*, 94–99.

- Roskos-Ewoldsen, D. R., & Fazio, R. H. (1992). On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology*, *63*, 198–211.
- Schutzwahl, A., & Borgstedt, K. (2005). The processing of affectively valenced stimuli: The role of surprise. *Cognition and Emotion*, *19*, 583–600.
- Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception and Psychophysics*, *49*, 83–90.
- Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception & Psychophysics*, *64*, 764–770.
- Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go here we want them to go: Capture of the eyes by new objects. *Psychological Science*, *9*, 379–385.
- Vuilleumier, P. (2002). Facial expression and selective attention. *Current Opinion in Psychiatry*, *15*(3), 291–300.
- Warrington, E. K., & James, M. (1967). An experimental investigation of facial recognition in patients with unilateral cerebral lesions. *Cortex*, *3*, 317–326.
- Yiend, J., & Mathews, A. (2001). Anxiety and attention to threatening pictures. *Quarterly Journal of Experimental Psychology*, *54A*, 665–681.

Manuscript received June 2005
Manuscript accepted October 2005