Effects of Gaze Cues on Object-Based Attention and the Underlying Mechanism

Authors: Yan Chi, Yunfei Gao, Hu Saisai, Song Fangxing, Wang Yonghui, Zhao Jingjing, Wang Yonghui, Zhao Jingjing

Date: 2022-01-31T19:16:09Z

Abstract

Eye gaze plays an important role in human social interaction and in capturing and maintaining attention. However, how eye gaze interacts with objects to guide attention allocation remains unclear to date. Therefore, this study adopted a two-rectangle cueing paradigm, using faces or objects with two gaze directions (direct gaze or averted gaze) as objects, and manipulated the SOA between cue and target to investigate the influence of eye gaze on object-based attention and its underlying mechanism. Experiment 1 found that the influence of eye gaze on object-based effects only occurred at the 300 ms SOA; further analysis revealed that the difference in object-based effects stemmed from participants responding faster to targets at invalid same-object locations under direct gaze conditions than under averted gaze conditions. This indicates that, compared to averted gaze, direct gaze better captures our attention, thereby generating larger object-based effects. Experiment 2 reversed the contrast of faces, ruling out the influence of low-level physical features on the results of Experiment 1. Experiment 3 used eyes with two gaze directions superimposed on cups as objects to explore whether the gaze direction effect could extend to other objects, and the results were the same as in Experiment 1. The results of this study indicate that eye gaze can interact with objects to jointly guide attention allocation, direct gaze better captures attention, but is influenced by SOA, and these findings support the sensory enhancement theory.

Full Text

The Effect of Gaze Direction on Object-Based Attention and Its Underlying Mechanism

YAN Chi[1], **GAO Yunfei**[1], **HU Saisai**[1], **SONG Fangxing**[2,3], **WANG Yonghui**[1*] (Corresponding Author), **ZHAO Jingjing**[1*] (Corresponding Author)

[1] School of Psychology, Shaanxi Normal University; Shaanxi Key Laboratory of Behavior and Cognitive Neuroscience; Shaanxi Provincial Key Research Center for Child and Adolescent Mental and Behavioral Health, Xi' an 710062, China
[2] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
[3] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract

Eye gaze plays a crucial role in human social interaction and in capturing and maintaining attention. However, how gaze direction interacts with objects to guide attentional allocation remains unclear. This study employed the two-rectangle cueing paradigm, using faces or objects with two gaze directions (direct or averted) as stimuli, and manipulated the stimulus onset asynchrony (SOA) between cue and target to investigate the effect of gaze direction on object-based attention and its underlying mechanism. Experiment 1 found that the influence of gaze direction on the object-based effect only emerged at the 300 ms SOA. Further analysis revealed that the difference in object-based effects stemmed from participants' faster responses to targets at invalid same-object locations under direct gaze compared to averted gaze conditions. This indicates that direct gaze captures attention more effectively than averted gaze, thereby producing a larger object-based effect. Experiment 2 reversed the contrast of the faces to rule out the influence of low-level physical features on the results from Experiment 1. Experiment 3 used objects consisting of eyes with two gaze directions superimposed on cups to explore whether the gaze effect could extend to other objects, and the results mirrored those of Experiment 1. The findings demonstrate that gaze direction can interact with objects to jointly guide attentional allocation, with direct gaze more effectively capturing attention, though this effect is modulated by SOA. These results support the sensory enhancement theory.

Keywords: object-based attention, gaze direction, social interaction, sensory enhancement theory, top-down processing

1 Introduction

Eye gaze serves a primary function in interpersonal communication and emotional connection (Kleinke & Chris, 1986). The ability to decode changes in gaze direction forms the foundation of social cognitive development (Farroni et al., 2002) and is essential for improving social interaction, regulating cognitive processing, and predicting future behavior (Bristow et al., 2007). Previous research has shown that gaze direction can guide attentional allocation (Chita-Tegmark, 2016; Fischer, 1999; Song et al., 2021). For instance, direct gaze can capture attention, with direct-gaze faces being identified faster and perceived as more attractive than averted-gaze faces (胡中华等, 2013; Böckler et al., 2014). Direct gaze can also maintain attention, as evidenced by difficulty disengaging from direct-gaze faces (Hietanen et al., 2016; Dalmaso et al., 2017). The eye region in direct gaze has a larger area of black-white contrast, which may lead to prolonged attentional dwell time (Senju & Hasegawa, 2005). Moreover, preferential attention to gaze emerges from infancy, with infants showing greater preference for open-eyed faces, particularly those with direct gaze (Farroni et al., 2002).

Objects can also guide attentional allocation (Chen, 2012; Lee & Shomstein, 2013; Shomstein & Behrmann, 2008). This object-guided attentional allocation is termed object-based attention (OBA), which refers to the phenomenon where individuals identify two (or more) features within the same object faster and more accurately than features across different objects (Shomstein, 2012). However, traditional spatial cueing paradigms cannot completely separate spatial attention from object-based attention. Egly et al. (1994) first dissociated OBA from space-based attention (SBA) using the two-rectangle cueing paradigm. In this paradigm, participants are initially presented with two horizontally or vertically arranged rectangles and a fixation point. A cue then appears at one end of either rectangle, followed by a target that appears randomly at the cued location (valid condition), at the opposite end of the cued rectangle (invalid same-object condition, ISO), or on the non-cued rectangle (invalid different-object condition, IDO). Critically, the ISO and IDO locations are equidistant from the cue location. Results show that participants respond significantly faster to targets at valid locations than at ISO and IDO locations, demonstrating a spatial attention effect. Additionally, participants respond significantly faster to targets at ISO locations than at IDO locations, revealing an object-guided attentional advantage effect.

In social interaction, beyond facial gaze information, the object properties of faces themselves also guide our attentional allocation. Research has found that attention to faces can activate the fusiform face area (FFA) and the inferior frontal junction (IFJ), with IFJ considered the neural basis of OBA (Baldauf & Desimone, 2014). Furthermore, as previously mentioned, objects without gaze information can themselves guide attentional allocation (Hu et al., 2020; Egly et al., 1994; Shomstein & Behrmann, 2008; Shomstein & Johnson, 2013; Song et al., 2020; Yeshurun & Rashal, 2017; Zhao et al., 2020). Therefore, an important question concerns how and whether social signals carried by gaze direction interact with objects to guide attention. Song et al. (2021) addressed this question using face and inverted face stimuli with a 300 ms SOA, finding that the OBA effect was significantly larger under direct gaze than averted gaze conditions, suggesting that gaze signals carrying social information can interact with face objects to guide attentional allocation. They obtained similar effects using eyes superimposed on rectangles, leading them to propose that gaze direction influences OBA through a general mechanism.

However, previous research on gaze direction's influence on OBA has several limitations. First, Song et al. (2021) used rectangles and real objects that differed in perceptual properties, complexity, and three-dimensional context, which cannot demonstrate that gaze direction' s influence on OBA has general applicability.

(1) Rectangles have simple, uniform properties, whereas real objects contain both top-down semantic properties and bottom-up low-level feature properties (Malcolm et al., 2015). Previous research has shown that objects' top-down and bottom-up properties differentially influence object-based attentional selection (Hu et al., 2020). (2) Rectangles are simple two-dimensional shapes that differ substantially from real objects in complexity. Research has found that attentional resources spread gradiently from the cue location, and region complexity can hinder this gradient spread, thereby affecting responses to targets (Chen et al., 2020). (3) In real life, we interact more with real three-dimensional objects in our environment than with two-dimensional shapes (Korisky & Mudrik, 2021). Real three-dimensional objects can also guide attentional allocation (Stephenson et al., 2017; Bayliss et al., 2006, 2007; Hudson et al., 2015) and are more likely to capture attention (Korisky & Mudrik, 2021). However, faces are more biological than other real three-dimensional objects (Crouzet et al., 2010) and possess unique processing advantages (Zhou et al., 2021) and attentional biases (Simpson et al., 2014; Langton et al., 2008). Whether the influence of gaze direction on OBA observed in face objects applies to other real objects remains unknown. Given that cups are common real objects frequently used by researchers to investigate joint attention (Stephenson et al., 2017; Bayliss et al., 2012) and preference evaluation (Bayliss et al., 2006, 2007), this study employed cups as stimuli in addition to faces to examine whether the effect of gaze direction on OBA persists after face processing is eliminated, and whether the underlying cognitive mechanisms are the same for other real, non-biological objects.

Second, stimulus onset asynchrony (SOA) between cue and target is an important factor influencing OBA (Jeurissen et al., 2016), yet no study has directly investigated the time course of gaze direction' s influence on OBA. On one hand, visual attention is a rhythmic processing mechanism (Chakravarthi et al., 2012; Landau & Fries, 2012; VanRullen, 2013), and as SOA increases, the OBA effect first increases then decreases. Previous gaze cueing studies using spatial cueing paradigms have found that significant cueing effects can be produced at short SOAs (105 ms) in detection tasks, with more stable effects at longer SOAs (300 ms), and disappearance of cueing effects at very long SOAs (1005 ms) (Friesen & Kingstone, 1998). Domestic researchers manipulating SOAs of 200 ms, 300 ms, and 500 ms found larger gaze cueing effects at 300 ms SOA (张智君等, 2015). Jeurissen et al. (2016) randomly presented SOAs ranging from 200–600 ms and found that the OBA effect at 300 ms was larger than at 200 ms and 600 ms. Additionally, previous OBA research has found that objects' physical features such as color influence OBA effects regardless of presentation duration, suggesting that bottom-up processing continuously affects OBA (Shomstein & Behrmann, 2008). Shomstein and Yantis (2004) found that at 400 ms SOA, OBA was modulated by top-down target probability based on individuals' search strategies, while at 600 ms SOA, the object effect disappeared as attention became completely guided by target probability, though at 200 ms SOA, OBA was not influenced by target probability.

On the other hand, gaze direction plays a crucial role in conveying social information about attentional direction and mental states (Frischen et al., 2007), involving top-down social cognitive processes such as gaze interpretation, empathy, and social attention (Kawai, 2011), which require longer processing time to take effect. Ristic and Kingstone (2005) manipulated SOAs of 100, 300, 600, and 1000 ms, presenting participants with an image that could be perceived as either a car or eyes. They found that when the stimulus was perceived as eyes, significant cueing effects emerged at 300 ms and 600 ms SOAs, but not when perceived as a car. This indicates that although gaze information captures attention in a bottom-up manner, it is influenced by top-down processing and requires time to become effective. Therefore, to further investigate whether the top-down information of gaze direction in Song et al.' s (2021) study already influences OBA effects at short SOAs and whether it continues to exert influence at long SOAs, this study manipulated SOAs of 100 ms, 300 ms, and 600 ms. We hypothesized that at 100 ms SOA, the OBA effect would not be influenced by gaze direction; at 300 ms SOA, gaze direction would affect the OBA effect, with the OBA effect being larger under direct gaze than averted gaze; but this influence would disappear at 600 ms SOA as the object effect diminishes.

In summary, this study used the two-rectangle cueing paradigm to investigate the influence of gaze direction on OBA under different SOA conditions. The study comprised three experiments. Experiment 1 used direct-gaze and averted-gaze faces as stimuli to examine how gaze direction interacts with face objects to influence selective attention. To exclude the influence of low-level physical features, Experiment 2 reversed the contrast of the faces (Ramamoorthy et al., 2019). Experiment 3 used cups superimposed with eyes to investigate whether, after face perception is eliminated, gaze direction can still influence OBA in real, non-biological objects. The hypothesis was that the object-based effect would be larger under direct gaze than averted gaze. If this difference stemmed from different response patterns to ISO location targets under direct versus averted gaze conditions, it would indicate that direct-gaze faces capture attention, leading to faster responses to targets at non-cued locations within direct-gaze objects. If the difference stemmed from different response patterns to IDO location targets, it would indicate that direct-gaze faces maintain attention, leading to slower responses to targets on averted-gaze objects. If differences emerged for both ISO and IDO locations, it would indicate that direct-gaze faces both capture and maintain attention.

2 Experiment 1: The Influence of Gaze Direction on Object-Based Attention in Faces

2.1.1 Participants

Based on a priori power analysis using G*Power, we assessed the required sample size for a medium effect size (f = 0.25; Cohen, 1992) for a 2×3×3 interaction, with = 0.05 and statistical power = 0.95 (Faul et al., 2009), which yielded a sample size of 14 participants. However, based on sample sizes used in previous OBA studies (Zhao et al., 2020; Yeshurun & Rashal, 2017), we determined a final planned sample size of 20-30 participants to ensure adequate statistical power. Twenty-five undergraduate students from Shaanxi Normal University participated in Experiment 1 (23 females, 2 males), aged 17-23 years ($M = 18$, $SD = 1.39$). All participants were right-handed with normal or corrected-to-normal vision and had no prior experience with similar experiments. Informed consent was obtained from all participants before the experiment, and they received compensation upon completion.

2.1.2 Materials and Apparatus

The experiment was programmed using E-Prime 2.0 and presented on a 60 Hz monitor. Participants viewed the display from a distance of approximately 63 cm. The stimuli consisted of two faces measuring $2° \times 4.9°$ of visual angle, with a central fixation point of $0.48° \times 0.48°$ positioned between them. The distance between the two faces was 0.9° of visual angle. The cue was a yellow outline that appeared randomly at either the top or bottom end of a face. The target was a white dot with a diameter of 0.48°. Participants rated the faces on a 9-point scale for pleasantness (1 = extremely unpleasant, 9 = extremely pleasant) and arousal (1 = extremely calm, 9 = extremely excited). The ratings showed no significant difference between direct-gaze ($M = 4.97$, $SE = 0.36$) and averted-gaze ($M = 5.27$, $SE = 0.40$) faces for pleasantness, $p > 0.05$, nor for arousal (direct: $M = 5.63$, $SE = 0.29$; averted: $M = 5.17$, $SE = 0.35$), $p > 0.05$.

2.1.3 Design and Procedure

The experiment employed a 2 (cue location: direct-gaze face, averted-gaze face) × 3 (cue validity: valid, invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms) within-subjects design. Cue location referred to whether the cue appeared on the direct-gaze or averted-gaze face, as shown in the cue display in Figure 1. All variables were within-subjects, with response time and accuracy as dependent variables. The formal experiment lasted approximately 60 minutes and consisted of 6 blocks with a total of 1,038 trials. To maintain participant alertness, the experiment included blank trials without targets. There were 864 target trials and 174 blank trials. In target trials, the probability of the target appearing at the cued location (valid) was 50%, while the probability of appearing at the opposite end of the cued rectangle (ISO) or on the non-cued rectangle (IDO) was 25% each.

The experimental procedure is illustrated in Figure 1. Each trial began with a 1000 ms presentation of a central fixation cross "+" and two faces. Subsequently, a yellow cue (RGB: 255, 255, 0) appeared at either end of the faces for 100 ms. After an additional 0 ms, 200 ms, or 500 ms interval, a target dot appeared at one of the four ends of the two faces (except in blank trials). The target remained on screen until participants responded or for 1500 ms if no response was made. A 500 ms blank screen preceded the next trial. Participants were instructed to press the "M" key as quickly as possible when they detected a target and to withhold responses on blank trials. Before the formal experiment, participants completed a practice block of 20 trials and were required to achieve at least 90% accuracy to proceed to the formal experiment.

Figure 1. Procedure for Experiment 1

2.2 Data Analysis and Results

Data were analyzed using SPSS 16.0. All participants achieved accuracy rates above 90%, with an average false alarm rate of 2.0% on blank trials. Prior to analysis, error trials and response times exceeding three standard deviations were excluded, resulting in a removal rate of 1.8%. A three-way repeated measures ANOVA was conducted on response times with 2 (cue location: direct-gaze, averted-gaze) × 3 (cue validity: valid, invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms).

The ANOVA revealed a significant main effect of SOA, $F(2, 48) = 78.38$, $p < 0.001$, p 2 = 0.77. Post-hoc tests showed that responses at 300 ms SOA ($M = 326$ ms, $SE = 8.67$) were faster than at 100 ms SOA ($M = 357$ ms, $SE = 8.74$), $p < 0.001$, 95% CI = [-36.35, -24.77], and at 600 ms SOA ($M = 350$ ms, $SE = 8.85$), $p < 0.001$, 95% CI = [-28.22, -18.07]. No significant difference was found between 100 ms and 600 ms SOA, $p = 0.094$, 95% CI = [-15.76, 0.93]. The main effect of cue location was not significant, $F(1, 24) = 0.87$, $p = 0.361$, p 2 = 0.04. The main effect of cue validity was significant, $F(2, 48) = 6.06$, $p = 0.009$, p 2 = 0.20, with significantly faster responses in the ISO condition ($M = 341$ ms, $SE = 8.77$) than in the IDO condition ($M = 348$ ms, $SE = 8.85$), indicating a significant OBA effect. No significant differences were found between the valid condition and ISO or IDO conditions. The interaction between SOA and cue location was not significant, $F(2, 48) = 1.83$, $p = 0.175$, p 2 = 0.07. The interaction between SOA and cue validity was significant, $F(4, 96) = 3.56$, $p = 0.013$, p 2 = 0.13. The interaction between cue validity and cue location was not significant, $F(2, 48) = 0.52$, $p = 0.592$, p 2 = 0.02. The three-way interaction between SOA, cue validity, and cue location was marginally significant, $F(4, 96) = 2.30$, $p = 0.081$, p 2 = 0.09.

Table 1. Mean response times (M ± SD) in Experiment 1 as a function of SOA, cue location, and cue validity

Cue Location	Cue Validity	100 ms	300 ms	600 ms
Direct Gaze	Valid	358 ± 43	321 ± 44	350 ± 47
	ISO	353 ± 47	320 ± 45	349 ± 47
	IDO	360 ± 43	333 ± 45	350 ± 44
Averted Gaze	Valid	359 ± 41	326 ± 41	350 ± 47
	ISO	351 ± 46	327 ± 47	346 ± 43
	IDO	361 ± 48	331 ± 46	352 ± 47

To further investigate the influence of SOA on OBA, we conducted a 2 (cue location: direct-gaze, averted-gaze) × 2 (cue validity: invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms) repeated measures ANOVA on response times. Results showed a significant main effect of SOA, F (2, 48) = 65.78, $p < 0.001$, p 2 = 0.73. The main effect of cue location was not significant, F (1, 24) = 0.11, p = 0.740, p 2 = 0.01. The main effect of cue validity was significant, F (1, 24) = 18.02, $p < 0.001$, p 2 = 0.43. All two-way interactions were non-significant (ps > 0.05). Importantly, the three-way interaction between SOA, cue validity, and cue location was significant, F (2, 48) = 3.79, p = 0.037, p 2 = 0.14.

We therefore conducted separate 2 (cue location: direct-gaze, averted-gaze) × 2 (cue validity: invalid same-object, invalid different-object) repeated measures ANOVAs for each SOA condition. At 100 ms SOA, the main effect of cue validity was significant, F (1, 24) = 36.28, $p < 0.001$, p 2 = 0.60, indicating that bottom-up attention to the cue produced an OBA effect. The main effect of cue location was not significant, F (1, 24) = 0.12, p = 0.733, p 2 = 0.01. The interaction between cue location and cue validity was not significant, F (1, 24) = 0.28, p = 0.603, p 2 = 0.01, suggesting that OBA effects did not differ between direct-gaze and averted-gaze conditions at 100 ms SOA.

At 300 ms SOA, the main effect of cue validity was significant, F (1, 24) = 8.98, p = 0.006, p 2 = 0.27, indicating an OBA effect. The main effect of cue location was not significant, F (1, 24) = 1.78, p = 0.195, p 2 = 0.07. Importantly, the interaction between cue location and cue validity was significant, F (1, 24) = 7.17, p = 0.013, p 2 = 0.23, indicating that top-down attention to the cue and top-down attention to gaze direction interactively influenced attentional allocation, resulting in different OBA effects between direct-gaze and averted-gaze conditions at 300 ms SOA. Further analysis revealed that the OBA effect was significantly larger in the direct-gaze condition (M = 12.79 ms, SE = 3.29) than in the averted-gaze condition (M = 3.94 ms, SE = 3.20), t (24) = -2.68, p = 0.013, Cohen' s d = 0.55, 95% CI = [-15.67, -2.03]. This difference arose from faster responses to ISO location targets in the direct-gaze condition compared to the averted-gaze condition, t(24) = 2.59, p = 0.016, Cohen' s d = 0.15, 95% CI = [1.40, 12.37]. No difference was found between direct-gaze and averted-gaze conditions for IDO location targets, t (24) = -0.86, p = 0.396, Cohen' s d = 0.04, 95% CI = [-6.67, 2.73]. These results indicate that direct gaze captures attention more effectively than averted gaze, supporting an attentional capture rather than attentional maintenance hypothesis.

At 600 ms SOA, neither main effects nor interactions were significant (ps > 0.05), indicating that no OBA effect emerged under either direct-gaze or averted-gaze conditions at this interval.

Figure 2. Results of Experiment 1

Note: Error bars represent standard errors of the mean. ** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$*

Analysis of accuracy rates revealed no significant main effects or interactions ($ps > 0.05$), indicating no speed-accuracy trade-off in this experiment.

The results of Experiment 1 demonstrated that at 300 ms SOA, the object-based effect was larger under direct gaze than averted gaze, with this difference arising from significantly faster detection of targets at ISO locations under direct gaze compared to averted gaze. This indicates that direct gaze captures attention. Additionally, no OBA effect was observed at 600 ms SOA, suggesting that the OBA effect decreases as the interval between cue and target increases, consistent with previous research (Chou & Yeh, 2018; Drummond & Shomstein, 2010; Shomstein & Yantis, 2004). To rule out the influence of low-level physical features such as eye whites and luminance contrast, Experiment 2 reversed the contrast of the faces. Research has shown that when the relative contrast between sclera and pupil changes from positive (iris darker than sclera) to negative polarity (sclera darker than iris), effects caused by eye whites and contrast remain unchanged, but adults become very inaccurate at judging gaze direction (Ricciardelli, 2009), and gaze effects disappear (Ramamoorthy et al., 2019). The hypothesis for Experiment 2 was that if no difference in object-based effects between direct-gaze and averted-gaze conditions emerged at 300 ms SOA, it would indicate that the results of Experiment 1 were due to encoding of gaze information rather than low-level physical features. If similar results to Experiment 1 were obtained, it would suggest that low-level features contributed to the observed effects.

3 Experiment 2: Contrast-Reversed Faces

3.1.1 Participants

Twenty-five undergraduate students from Shaanxi Normal University were recruited for this experiment (23 females, 2 males), aged 17–21 years ($M = 19$, $SD = 1.03$). All participants had normal or corrected-to-normal vision and had never participated in similar experiments. Informed consent was obtained before the experiment, and participants received compensation upon completion.

3.1.2 Materials and Apparatus

The materials and apparatus were similar to Experiment 1, with the exception that the stimuli were contrast-reversed versions of the faces used in Experiment 1. All other aspects remained identical to Experiment 1.

3.1.3 Design and Procedure

The experimental design and procedure were identical to Experiment 1, as illustrated in Figure 3.

Figure 3. Procedure for Experiment 2

3.2 Data Analysis and Results

All participants achieved accuracy rates above 90%, with an average false alarm rate of 1.6% on blank trials. The data analysis method was identical to Experiment 1. Prior to analysis, error trials and response times exceeding three standard deviations were excluded, resulting in a removal rate of 2.2%. A three-way repeated measures ANOVA was conducted with 2 (cue location: direct-gaze, averted-gaze) × 3 (cue validity: valid, invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms).

The ANOVA revealed a significant main effect of SOA, $F (2, 48) = 90.95$, $p < 0.001$, $p\ 2 = 0.79$. Post-hoc tests showed that responses at 300 ms SOA (M = 343 ms, SE = 9.09) were faster than at 100 ms SOA (M = 378 ms, SE = 8.95), $p < 0.001$, 95% CI = [-40.41, -29.36], and at 600 ms SOA (M = 369 ms, SE = 9.24), $p < 0.001$, 95% CI = [-33.58, -19.30]. Responses at 600 ms SOA were also faster than at 100 ms SOA, $p = 0.035$, 95% CI = [-16.39, -0.50]. The main effect of cue location was significant, $F (1, 24) = 11.93$, $p = 0.002$, $p\ 2 = 0.33$, with significantly faster responses when the cue appeared on averted-gaze faces (M = 362 ms, SE = 8.93) than on direct-gaze faces (M = 365 ms, SE = 9.00). The main effect of cue validity was significant, $F (2, 48) = 16.17$, $p < 0.001$, $p\ 2 = 0.40$, with significantly faster responses in the valid condition (M = 361 ms, SE = 8.71) than in the IDO condition, and faster responses in the ISO condition (M = 360 ms, SE = 8.79) than in the IDO condition (M = 369 ms, SE = 9.49), indicating a significant OBA effect. The interaction between SOA and cue validity was significant, $F (4, 96) = 4.92$, $p = 0.003$, $p\ 2 = 0.17$. The interaction between SOA and cue location was significant, $F (2, 48) = 5.76$, $p = 0.006$, $p\ 2 = 0.19$. The interaction between cue validity and cue location was not significant, $F (2, 48) = 1.44$, $p = 0.247$, $p\ 2 = 0.06$. The three-way interaction between SOA, cue validity, and cue location was not significant, $F (4, 96) = 2.20$, $p = 0.095$, $p\ 2 = 0.08$.

Table 2. Mean response times (M ± SD) in Experiment 2 as a function of SOA, cue location, and cue validity

Cue Location	Cue Validity	100 ms	300 ms	600 ms
Direct Gaze	Valid	382 ± 47	339 ± 46	368 ± 46
	ISO	372 ± 42	340 ± 42	368 ± 47
	IDO	388 ± 49	354 ± 54	369 ± 47
Averted Gaze	Valid	375 ± 44	335 ± 41	368 ± 46
	ISO	371 ± 47	341 ± 48	368 ± 47
	IDO	378 ± 46	347 ± 49	375 ± 50

To further investigate the influence of SOA on OBA, we conducted a 2 (cue location: direct-gaze, averted-gaze) × 2 (cue validity: invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms) repeated measures ANOVA. Results showed a significant main effect of SOA, $F (2, 48) = 64.02$, p

< 0.001, $p\ 2 = 0.73$. The main effect of cue location was marginally significant, $F\ (1, 24) = 3.96$, $p = 0.058$, $p\ 2 = 0.14$. The main effect of cue validity was significant, $F\ (1, 24) = 31.64$, $p < 0.001$, $p\ 2 = 0.57$. The interaction between SOA and cue location was significant, $F\ (2, 48) = 5.66$, $p = 0.008$, $p\ 2 = 0.19$. All other two-way interactions were non-significant (ps > 0.05). Importantly, the three-way interaction between SOA, cue validity, and cue location was significant, $F\ (2, 48) = 3.34$, $p = 0.046$, $p\ 2 = 0.12$, indicating that OBA was influenced by both cue validity and cue location.

We therefore conducted separate 2 (cue location: direct-gaze, averted-gaze) $\times$ 2 (cue validity: invalid same-object, invalid different-object) repeated measures ANOVAs for each SOA condition. At 100 ms SOA, the main effect of cue validity was significant, $F\ (1, 24) = 22.08$, $p < 0.001$, $p\ 2 = 0.48$, indicating that bottom-up attention to the cue produced an OBA effect. The main effect of cue location was significant, $F\ (1, 24) = 10.12$, $p = 0.004$, $p\ 2 = 0.30$, indicating differential processing of cues on direct-gaze versus averted-gaze faces. The interaction between cue location and cue validity was not significant, $F\ (1, 24) = 3.28$, $p = 0.083$, $p\ 2 = 0.12$, suggesting that OBA effects did not differ between direct-gaze and averted-gaze conditions at 100 ms SOA.

At 300 ms SOA, the main effect of cue validity was significant, $F\ (1, 24) = 12.68$, $p = 0.002$, $p\ 2 = 0.35$, indicating an OBA effect. The main effect of cue location was marginally significant, $F\ (1, 24) = 3.70$, $p = 0.066$, $p\ 2 = 0.13$. The interaction between cue location and cue validity was not significant, $F\ (1, 24) = 2.36$, $p = 0.138$, $p\ 2 = 0.09$, indicating that OBA effects did not differ between direct-gaze and averted-gaze conditions at 300 ms SOA.

At 600 ms SOA, neither main effects nor interactions were significant (ps > 0.05), indicating that no OBA effect emerged under either direct-gaze or averted-gaze conditions at this interval.

Figure 4. Results of Experiment 2

Note: Error bars represent standard errors of the mean. ** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$*

Analysis of accuracy rates revealed no significant main effects or interactions (ps > 0.05), indicating no speed-accuracy trade-off in this experiment.

Experiment 2 only reversed the contrast of the faces from Experiment 1 without changing the contrast differences. The absence of a difference in object-based effects between direct-gaze and averted-gaze conditions at 300 ms SOA demonstrates that the results of Experiment 1 were due to encoding of gaze information rather than low-level physical features such as contrast. The lack of an OBA effect at 600 ms SOA was consistent with Experiment 1. To examine whether the influence of gaze direction on OBA could extend to other real objects after face perception was eliminated, Experiment 3 used non-biological cups superimposed with eyes as stimuli. If no gaze effect emerged in Experiment 3, it would suggest that gaze direction' s influence on OBA depends on processing of the entire face. If results similar to Experiment 1 were obtained, it would indicate that gaze direction' s influence on attentional allocation can extend to other real objects after face information is eliminated.

4 Experiment 3: Eyes Superimposed on Cups

4.1.1 Participants

Twenty-five undergraduate students from a university were recruited for this experiment (23 females, 2 males), aged 17–20 years (M = 18, SD = 0.64). All participants had normal or corrected-to-normal vision and had never participated in similar experiments. Informed consent was obtained before the experiment, and participants received compensation upon completion.

4.1.2 Materials and Apparatus

The materials and apparatus were similar to Experiment 1, with the exception that the stimuli consisted of two grayscale cups superimposed with eyes. All other aspects remained identical to Experiment 1.

4.1.3 Design and Procedure

The experimental design and procedure were identical to Experiment 1, as illustrated in Figure 5.

Figure 5. Procedure for Experiment 3

4.2 Data Analysis and Results

All participants achieved accuracy rates above 90%, with a false alarm rate of 1.7% on blank trials. The data analysis method was identical to Experiment 1. Prior to analysis, error trials and response times exceeding three standard deviations were excluded, resulting in a removal rate of 2.3%. A three-way repeated measures ANOVA was conducted with 2 (cue location: direct-gaze, averted-gaze) × 3 (cue validity: valid, invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms).

The ANOVA revealed a significant main effect of SOA, $F(2, 48) = 48.29$, $p < 0.001$, p 2 = 0.67. Post-hoc tests showed that responses at 300 ms SOA (M = 357 ms, SE = 9.48) were faster than at 100 ms SOA (M = 391 ms, SE = 9.84), $p < 0.001$, 95% CI = [-41.58, -27.17], and at 600 ms SOA (M = 382 ms, SE = 7.71), $p < 0.001$, 95% CI = [-33.60, -16.72]. No significant difference was found between 100 ms and 600 ms SOA, $p = 0.163$, 95% CI = [-20.93, 2.51]. The main effect of cue location was not significant, $F(1, 24) = 0.39$, $p = 0.536$, p 2 = 0.02. The main effect of cue validity was significant, $F(2, 48) = 11.73$, $p < 0.001$, p 2 = 0.33, with significantly faster responses in the valid condition (M = 372 ms, SE = 8.13) than in the IDO condition, and faster responses in the ISO condition (M = 375 ms, SE = 9.36) than in the IDO condition (M = 382 ms, SE = 9.16), indicating a significant OBA effect. The interaction between SOA and cue validity was significant, $F(4, 96) = 3.58$, $p = 0.015$, p 2 = 0.13. The interaction between SOA and cue location was not significant, $F(2, 48) = 0.93$, $p = 0.395$, p 2 = 0.04. The interaction between cue validity and cue location was not significant, $F(2, 48) = 2.06$, $p = 0.138$, p 2 = 0.08. The three-way interaction between SOA, cue validity, and cue location was marginally significant, $F(4, 96) = 2.45$, $p = 0.063$, p 2 = 0.09.

Table 3. Mean response times (M ± SD) in Experiment 3 as a function of SOA, cue location, and cue validity

Cue Location	Cue Validity	100 ms	300 ms	600 ms
Direct Gaze	Valid	389 ± 46	348 ± 38	381 ± 41
	ISO	388 ± 54	353 ± 51	382 ± 42
	IDO	400 ± 51	367 ± 51	384 ± 42
Averted Gaze	Valid	387 ± 44	349 ± 46	378 ± 39
	ISO	388 ± 53	363 ± 52	380 ± 39
	IDO	395 ± 52	361 ± 55	386 ± 39

To further investigate the influence of SOA on OBA, we conducted a 2 (cue location: direct-gaze, averted-gaze) × 2 (cue validity: invalid same-object, invalid different-object) × 3 (SOA: 100 ms, 300 ms, 600 ms) repeated measures ANOVA on response times. Results showed a significant main effect of SOA, $F(2, 48) = 33.62$, $p < 0.001$, p 2 = 0.58. The main effect of cue location was not significant, $F(1, 24) = 0.02$, $p = 0.893$, p 2 = 0.00. The main effect of cue validity was significant, $F(1, 24) = 17.82$, $p < 0.001$, p 2 = 0.43. All two-way interactions were non-significant (ps > 0.05). Importantly, the three-way interaction between SOA, cue validity, and cue location was significant, $F(2, 48) = 4.05$, $p = 0.025$, p 2 = 0.14.

We therefore conducted separate 2 (cue location: direct-gaze, averted-gaze) × 2 (cue validity: invalid same-object, invalid different-object) repeated measures ANOVAs for each SOA condition. At 100 ms SOA, the main effect of cue validity was significant, $F(1, 24) = 17.03$, $p < 0.001$, p 2 = 0.42, indicating that bottom-up attention to the cue produced an OBA effect. The main effect of cue location was not significant, $F(1, 24) = 1.81$, $p = 0.192$, p 2 = 0.07. The interaction between cue location and cue validity was not significant, $F(1, 24) = 0.59$, $p = 0.450$, p 2 = 0.02, suggesting that OBA effects did not differ between direct-gaze and averted-gaze conditions at 100 ms SOA.

At 300 ms SOA, the main effect of cue validity was significant, $F(1, 24) = 9.62$, $p = 0.005$, p 2 = 0.29, indicating an OBA effect. The main effect of cue location was not significant, $F(1, 24) = 0.41$, $p = 0.531$, p 2 = 0.02. Importantly, the interaction between cue location and cue validity was significant, $F(1, 24) = 8.13$, $p = 0.009$, p 2 = 0.25, indicating that top-down attention to the cue and top-down attention to gaze direction interactively influenced attentional allocation, resulting in different OBA effects between direct-gaze and averted-gaze conditions at 300 ms SOA. Further analysis revealed that the OBA effect was significantly larger in the direct-gaze condition ($M = 14.61$ ms, $SE = 3.41$) than in the averted-gaze condition ($M = -1.50$ ms, $SE = 3.64$), $t(24) = -2.85$, $p = 0.009$, Cohen' s $d = 0.91$, 95% CI = [-27.76, -4.45]. This difference arose from faster responses to ISO location targets in the direct-gaze condition compared to the averted-gaze condition, $t(24) = 2.35$, $p = 0.027$, Cohen' s $d = 0.19$, 95% CI = [1.23, 18.76]. No difference was found between direct-gaze and averted-gaze conditions for IDO location targets, $t(24) = -1.50$, $p = 0.147$, Cohen' s $d = 0.12$, 95% CI = [-14.51, 2.30]. These results indicate that direct gaze captures attention more effectively than averted gaze, supporting an attentional capture rather than attentional maintenance hypothesis.

At 600 ms SOA, neither main effects nor interactions were significant ($ps > 0.05$), indicating that no OBA effect emerged under either direct-gaze or averted-gaze conditions at this interval.

Figure 6. Results of Experiment 3

Note: Error bars represent standard errors of the mean. ** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$*

Analysis of accuracy rates revealed no significant main effects or interactions ($ps > 0.05$), indicating no speed-accuracy trade-off in this experiment.

The results of Experiment 3 mirrored those of Experiment 1, showing that at 300 ms SOA, the object-based effect was significantly larger under direct gaze than averted gaze, with this difference arising from different response patterns to ISO location targets between the two conditions. This indicates that direct gaze captures attention and further demonstrates that the influence of gaze direction on OBA can extend to other real objects. No OBA effect was observed at 600 ms SOA.

5 General Discussion

Through three experiments using faces and cups with gaze direction as stimuli, this study reveals that in the two-rectangle cueing paradigm, the influence of gaze direction on OBA is modulated by SOA. The effect of gaze direction on OBA emerges at longer SOAs (300 ms) and disappears at shorter (100 ms) and longer SOAs (600 ms). Experiment 1 used face stimuli with different SOAs to investigate whether and how gaze direction interacts with face objects to guide attentional allocation. At 300 ms SOA, participants detected targets at ISO locations faster under direct gaze than averted gaze, resulting in a larger OBA effect. This suggests that direct gaze captures attention more effectively than averted gaze. Experiment 2 reversed face contrast and found no difference in object-based effects between direct-gaze and averted-gaze conditions, indicating that the results of Experiment 1 were due to encoding of gaze information rather than low-level features such as contrast. Experiment 3 used cups superimposed with eyes and found similar results to Experiment 1 after face perception was eliminated. The three experiments consistently demonstrate that direct gaze captures attention, facilitating the allocation of attentional resources from the cued location to the ISO location. Moreover, gaze direction can influence object-based attentional allocation, and this influence extends to real objects.

The results show that object-based attention effects differ between direct-gaze and averted-gaze conditions, with this difference arising from faster detection of targets at ISO locations under direct gaze. This finding aligns with the sensory enhancement hypothesis of object-based attention. Currently, three main theories explain OBA effects: sensory enhancement theory (enhancement spreading theory), attentional shifting theory, and attentional priorization theory. Early sensory enhancement theory, proposed by Desimone and Duncan (1995) as the biased competition model, suggests that sensory enhancement may result from biased competition among neural representations of multiple objects, making the representation of cued objects more effective than that of uncued objects, thereby enabling prioritized selection. Selection of the cued object in turn facilitates faster and more accurate processing of features within that object. Current sensory enhancement theory posits that attention spreads within objects and is constrained by object boundaries, leading to enhanced sensory representations of cued objects and consequently faster and more accurate responses to features or items within those objects (Conci & Muller, 2009; Richard et al., 2008; Shomstein, 2012). In contrast, attentional shifting theory suggests that higher costs are incurred when shifting attention from one object to another (Brown & Denney, 2007; Ushitani et al., 2010; Yeshurun & Rashal, 2017). Attentional priorization theory proposes that OBA effects result from different attentional priorities at different locations during visual search, with visual search defaulting to locations within the cued object, leading to later search of non-cued objects (Drummond & Shomstein, 2010; Shomstein, 2012; Yeshurun & Rashal, 2017). If direct-gaze faces capture attention, they would be allocated more attentional resources, further enhancing their representational strength. Consequently, participants would detect targets at ISO locations faster under direct gaze than averted gaze, producing a larger OBA effect. This hypothesis aligns with sensory enhancement theory. If direct-gaze faces maintain attention, making disengagement from them more difficult than from averted-gaze faces, participants would detect targets at IDO locations slower under direct gaze than averted gaze, also producing a larger OBA effect. This hypothesis aligns with attentional shifting theory. If direct-gaze faces both capture and maintain attention, differences would emerge for both ISO and IDO locations. This hypothesis aligns with attentional priorization theory. Our findings show that direct gaze preferentially captured participants' attention, further enhancing its sensory representation and leading to better performance at ISO locations at 300 ms SOA. Thus, the OBA effect originated from top-down facilitatory processing of ISO locations by gaze direction, supporting the sensory enhancement theory.

The results also demonstrate that the OBA effect decreases as the interval between cue and target increases, consistent with previous research (Chou & Yeh, 2018; Drummond & Shomstein, 2010; Shomstein & Yantis, 2004). More importantly, the difference in object-based effects between direct-gaze and averted-gaze conditions only emerged at 300 ms SOA, consistent with previous findings (Song et al., 2021). This pattern can be explained as follows: when a cue appears suddenly, attention is captured bottom-up and shifts to the cued location (Posner, 1980). In the short term, only the cue exerts influence, which is why an OBA effect emerged at 100 ms SOA in our study, unaffected by gaze direction. As time progresses, gaze cues begin to take effect, consistent with previous findings that gaze direction is influenced by top-down processing and requires time to encode (Bayliss & Tipper, 2006; Kawai, 2011). Therefore, at 300 ms, an interaction between gaze direction and object-based attention emerged. With increasing SOA, participants may have developed preparatory strategies that reduced the importance of spatial orienting to the cue (彭姓等, 2019). Additionally, Lou and Lorist et al. (2021) found that object effects only emerged at specific cue-target intervals (130–190 ms, 230–260 ms, 500–530 ms). Thus, at 600 ms SOA, as the cueing effect disappeared, gaze direction no longer influenced OBA. In summary, our findings indicate that the influence of gaze direction on OBA is relatively brief, emerging late and disappearing quickly.

Notably, after face information was eliminated, the influence of gaze direction on OBA could still occur in other real objects, indicating that this influence has general applicability. This is consistent with previous research showing that isolated eyes or eyes superimposed on other stimuli can produce cueing effects after face disappearance (Green et al., 2013; Ristic & Kingstone, 2005). Itier et al.' s neural model of face processing proposes that two types of neurons may exist in the superior temporal sulcus (STS) region. When upright faces are present, face-selective neurons primarily function. However, when face contrast is reversed or faces are inverted, placing eyes in an abnormal face context, both eye-selective neurons and face neurons function together, producing similar and even larger N170 amplitudes than upright faces. This model attributes these results to processing of the eye region. In that study, amplitudes produced by isolated eye stimuli were always larger than those produced by faces, while amplitudes produced by faces without eyes were similar to those produced by faces, suggesting that the presence of eyes makes faces special (Itier et al., 2007). This aligns with our findings.

This study is the first to examine how OBA changes with SOA during social interaction. However, several limitations should be noted. First, the study focused primarily on gaze direction, but other factors related to faces, such as eye movements (Böckler et al., 2014; 陈艾睿等, 2014), face inversion (Murphy & Cook, 2017), and head orientation (Yumiko et al., 2018), may also influence OBA and require further investigation. Second, the gender ratio in our study was relatively unbalanced due to random recruitment. Research has shown that females have advantages over males in face recognition and processing social cues (Bayliss et al., 2005), so future studies should better control gender ratios. Additionally, debate continues regarding whether processing of social gaze cues differs from processing of non-social arrow cues (纪皓月等, 2017). Social attentional cues are more reflexive and less susceptible to top-down influence than non-social cues such as arrows (Friesen et al., 2004). Future research could explore whether arrow cues produce similar cognitive mechanisms to those observed in this study. Finally, the influence of gaze direction on OBA across different cultures warrants investigation (霍鹏辉等, 2021). For example, Japanese people consider extensive eye contact disrespectful, whereas in China it signals politeness and interest in the conversation (Shota et al., 2015). Asians typically show concentrated fixation on the central region of faces, while Caucasians primarily fixate on the eye and mouth regions (Blais et al., 2008).

6 Conclusion

1. Gaze direction can interact with objects to guide attentional allocation. Direct gaze more effectively captures attention, and top-down facilitatory processing of ISO locations by gaze direction leads to larger OBA effects, supporting the sensory enhancement theory.
2. The influence of gaze direction on OBA is modulated by SOA. This influence emerges relatively late and disappears quickly.
3. The influence of gaze direction on OBA has general applicability, occurring not only when faces serve as objects but also extending to other real objects.

References

The reference section appears to be already partially translated. The full reference list would be completed here following standard academic formatting.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.