

# SCIENTIFIC REPORTS



OPEN

## Human cortical activity evoked by contextual processing in attentional orienting

Shuo Zhao<sup>1,2,3</sup>, Chunlin Li<sup>4</sup>, Shota Uono<sup>5</sup>, Sayaka Yoshimura<sup>5</sup> & Motomi Toichi<sup>1,3</sup>

The ability to assess another person's direction of attention is paramount in social communication, many studies have reported a similar pattern between gaze and arrow cues in attention orienting. Neuroimaging research has also demonstrated no qualitative differences in attention to gaze and arrow cues. However, these studies were implemented under simple experiment conditions. Researchers have highlighted the importance of contextual processing (i.e., the semantic congruence between cue and target) in attentional orienting, showing that attentional orienting by social gaze or arrow cues could be modulated through contextual processing. Here, we examine the neural activity of attentional orienting by gaze and arrow cues in response to contextual processing using functional magnetic resonance imaging. The results demonstrated that the influence of neural activity through contextual processing to attentional orienting occurred under invalid conditions (when the cue and target were incongruent versus congruent) in the ventral frontoparietal network, although we did not identify any differences in the neural substrates of attentional orienting in contextual processing between gaze and arrow cues. These results support behavioural data of attentional orienting modulated by contextual processing based on the neurocognitive architecture.

The ability to assess another person's direction of attention is paramount in social communication. For example, we are able to identify a person's focus based on their eye gaze, thus enabling an understanding of other people's inner state (such as thoughts, beliefs, and desires)<sup>1</sup>. Similar to eye gaze, non-social stimuli also play important roles in influencing attention, such as an arrow on a road sign. However, compared with eye gazes, non-social directional stimuli are not helpful when making conclusions regarding someone's cognitive state, such as speculating about what a person wishes to do.

Over the past two decades, cognitive psychologists have focused on comparing the role of directional gaze and arrow cues in attentional orienting. These studies have typically investigated attentional orienting based on gaze and arrow cues using a modified version of Posner's cueing paradigm<sup>2</sup>. For example, Friesen and Kingstone<sup>3</sup> presented non-predictive gaze cues at the centre of a screen prior to the presentation of a peripheral target (right or left). Before the onset of the target, a centrally presented directional cue (e.g., eye gaze) appears on screen. Under valid conditions, the cue will accurately indicate the subsequent target location, whereas under invalid conditions, the cue will indicate the opposite location. A rapid response to a validly cued target indicates an allocation of attention (i.e., orienting) to the target location prior to target onset. In contrast, a delayed response to an invalidly cued target occurs when the onset of the target at the opposite location, indicating a reorienting of attention to the target. Previous studies<sup>4–9</sup> have commonly demonstrated that arrow cues automatically trigger attentional shifts in the same manner as gaze cues. These studies have demonstrated that both gaze and arrow cues trigger attentional shifts when they are counterpredictive of a target location<sup>7</sup>, facilitate response time when discriminating the target following the cue<sup>6</sup>, have comparable sensitivity to object-based selection<sup>4,5,8</sup> and the stimulus onset asynchrony between the cue and target<sup>9</sup>.

<sup>1</sup>Faculty of Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, 606-8507, Japan.

<sup>2</sup>International Research Fellow of the Japan Society for the Promotion of Science, Tokyo, 102-0083, Japan.

<sup>3</sup>Organization for Promoting Neurodevelopmental Disorder Research, Kyoto, 606-8392, Japan. <sup>4</sup>School of Biomedical Engineering, Capital Medical University, Beijing, 100069, China. <sup>5</sup>Department of Neurodevelopmental Psychiatry, Habilitation and Rehabilitation, Graduate School of Medicine, Kyoto University, Kyoto, 606-8507, Japan. Correspondence and requests for materials should be addressed to C.L. (email: [lichunlin1981@163.com](mailto:lichunlin1981@163.com)) or M.T. (email: [toichi.motomi.4v@kyoto-u.ac.jp](mailto:toichi.motomi.4v@kyoto-u.ac.jp))

Recent neuroimaging studies regarding attentional orienting have attempted to investigate differences in cortical activity between gaze and arrow cues. These studies have focused on two attentional networks (reviewed by<sup>10,11</sup>). The dorsal frontoparietal network, with regions centred around the intraparietal sulcus (IPS), superior parietal lobule (SPL)/Brodmann's area (BA)5, 7, and frontal eye field (FEF)/BA8, may be responsible for orienting of attention to a validly cued target in the cueing paradigm, but also for reorienting attention to an invalidly cued target<sup>10,11</sup>. The ventral frontoparietal network, with regions centred on the temporoparietal junction (TPJ)/BA39, 40, 22 and ventral frontal cortex (VFC)/BA44, 45, 47 (including parts of the middle frontal gyrus (MFG) and inferior frontal gyrus (IFG)), may only be responsible for reorienting attention. Most previous studies<sup>12–18</sup> have demonstrated that the differences in cortical activity associated with social gaze and arrow cues are quantitative rather than qualitative, although some studies<sup>19–21</sup> have reported evidence suggesting different mechanisms for these cues. For example, Tipper *et al.*<sup>17</sup> reported attentional orienting to both eye gaze and arrow cues engaged extensive dorsal and ventral frontoparietal networks, but the magnitude of activation differed between these networks. However, these studies only examined the differences between gaze and arrow cues under simple conditions (e.g. a dot or letter as the target). Given that Birmingham and Kingstone<sup>22</sup> suggested that the apparent difference in attentional orienting between gaze and arrow cues might be distinguished only when the cues were embedded in a rich environment, it is thus important to examine the differences between gaze and arrow cues under more complex conditions.

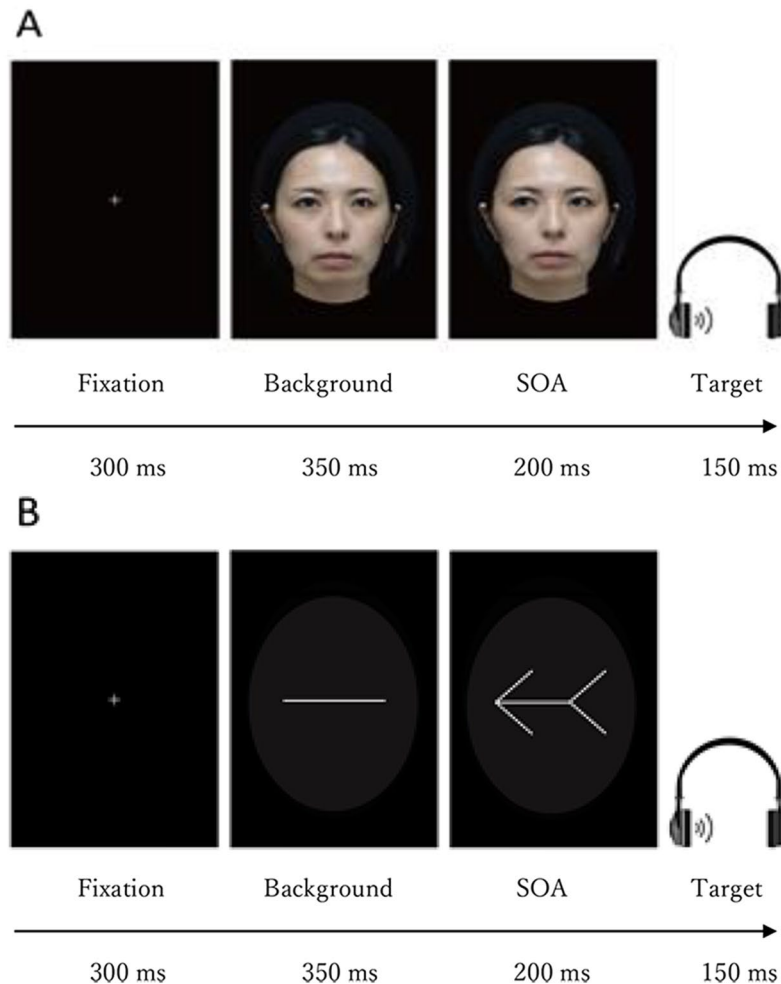
Some studies have highlighted the importance of contextual processing (i.e., the semantic congruence between the cue and target) in attentional orienting when using arrows or eye gaze as cues. Previous studies<sup>23,24</sup> have demonstrated that attentional orienting is facilitated through contextual processing when using arrows as cues. For example, Ristic *et al.*<sup>23</sup> examined attentional orienting based on whether facial gaze and arrow cues could be triggered through the contextual processing of cue-target colour contingencies. The results indicated that attentional orienting elicited by an arrow rather than an eye gaze was sensitive to colour-congruent target stimuli; an attentional orienting effect for blue arrows was only evident for blue targets. However, other studies<sup>25–27</sup> have demonstrated that attentional orienting with facial gaze was facilitated through a strongly contextual relationship between the cue and target when there was congruence in meaning between the cue and target. For example, Bayliss *et al.*<sup>25</sup> reported that compared with disgusted faces, the gaze direction of happy faces more effectively oriented attention to pleasant targets. These findings indicated that participants could employ contextual information in attentional orienting by arrows or eye gaze cues to effectively capture important information, although the context effect might be observed only when targets are presented at a specific context of colour and emotion for gazes and arrows. These findings raised a question regarding whether attentional orienting differs between eye gaze and arrow cues when these cues were influenced through contextual processing.

At a neural level, researchers have shown activity in the TPJ and superior temporal sulcus (STS)/BA21, 22 associated with contextual processing in attention. Geng and Vossel<sup>28</sup> reviewed previous evidence, indicating that the TPJ (anatomically, the TPJ is strictly defined as the cortex at the intersection of the posterior superior temporal, supramarginal, and angular gyri) was engaged in terms of “contextual updating” in attention. For example, Weidner *et al.*<sup>29</sup> demonstrated that cortical activity in TPJ increased when the contextual processing of the relationship between the cue and target was incongruent as opposed to congruent (i.e., when the target-defining dimension (orientation or colour) was incongruently rather than congruently cued). Moreover, Noppeney *et al.*<sup>30</sup> observed that the activity in STS increased when a sound or speech target was incongruent (e.g., a car picture paired with the spoken word ‘owl’) as opposed to congruent (e.g., a cat picture paired with the spoken word ‘cat’) with prior visual information. This finding indicated that context also modulated the activity of STS. Consistently, when a strong relationship was established between the target word and a word sound that had been previously presented, the results showed the enhanced activation for thematically related categories (e.g., picture + frame) and response suppression for taxonomically related categories (e.g., chair + armchair) in the left STS<sup>31</sup>. In the present study, we focused on these brain regions to examine the influence of neural systems in relation to gaze and arrow cues through contextual processing focused on the relationship between the cue and target in attentional orienting.

In the present study, we examined the neural activity of attentional orienting with social gaze and arrows as cues using Posner's cueing paradigm. Based on a previous study<sup>27</sup>, two sounds (a social voice and a tone) were manipulated as targets to determine the contextual relationship between cue and target; that is, social gaze and social voice and arrow and tone as congruent meaning conditions, and social gaze and tone and arrow and social voice as incongruent meaning conditions. The aims of this study are as follows: (1) We first wanted to examine whether the influence of neural activity in TPJ and STS differed between gaze and arrow cues in response to contextual processing of the relationship between cue and target. (2) Furthermore, given that previous studies<sup>10,11</sup> have characterised the functional mechanisms of the orienting and reorienting of attention (i.e., valid and invalid conditions) in dorsal and ventral frontoparietal networks, respectively, we considered it important to specifically investigate these functional mechanisms and how they were modulated through the contextual processing of the relationship between cue and target. Specifically, if different neural activity was observed for gaze and arrow in response to the contextual processing of cue-target, we would subsequently examine whether the neural activity for attentional orienting differed with contextual processing between gaze and arrow cues at the dorsal and ventral frontoparietal networks, respectively. In contrast, if no difference between gaze and arrow was evident, we would then examine only the influence of neural activity for attentional orienting by contextual processing in both gaze and arrow cues in these two attentional networks.

## Methods

**Participants.** This research was approved by the local ethics committee of Capital Medical University, Beijing, China. No foreseeable risk to the participants was present, and personal identifying information was not collected. Participants provided written informed consent and background information. All procedures complied



**Figure 1.** Illustration of the stimulus presentation.

with the ethical standards of the 1964 Declaration of Helsinki regarding the treatment of human participants in research. In total, 22 volunteers (9 women, 13 men; mean  $\pm$  SD age,  $22.95 \pm 2.61$  years) participated. All participants were right-handed, as assessed by the Edinburgh Handedness Inventory<sup>32</sup>, and had normal or corrected-to-normal visual and auditory acuity.

**Stimuli.** Visual and auditory stimuli were almost identical to those used in a previous behavioural study (at a sound level comfortable to each participant)<sup>27</sup>. Previous studies have demonstrated that female faces are less resemblance to angry expressions than male faces, and male faces are perceived as less likeable<sup>33,34</sup> and more powerful<sup>35</sup>. To avoid any differential influence of expression (e.g., anger), a Japanese female face with neutral expressions was used for this task (Fig. 1A). The image was obtained from a previous study<sup>36</sup>, in which the emotional intensity of facial stimuli with neutral expressions was assessed. The results confirmed that these facial images were considered neutral rather than emotional. Based on these findings, it is reasonable to propose that the female face image in the present study conveyed neutral facial expressions. Moreover, three versions of each face were produced: one version with a direction of gaze straight ahead, another version with the pupils averted leftward, and a third version with the pupils averted rightward. The faces measured approximately  $4.7^\circ$  wide and  $6.9^\circ$  high. For the arrow cue, a symmetrical arrow was presented as the cue stimulus, with an arrowhead at one end and a tail at the opposite end. The arrows measured  $4.7^\circ$  in width by  $1.7^\circ$  in height and were light grey.

Furthermore, two types of auditory stimuli were presented as targets. One type was sampled from a woman: an/i/voice sound (F0 frequency of 300 Hz), which is similar to the /iy/sound in English. The other type was a pure tone of similar frequency to the F0 voice (300 Hz), which was produced using the Audacity software package (ver. 1.3.13; Audacity store.com). The duration of the stimulus presentation was 150 ms.

**Apparatus.** These stimuli were generated on a computer and presented to the participants via a custom-built, magnet-compatible audio-visual system during magnetic resonance (MR) scanning. To attenuate the acoustic noise that accompanies fMRI (functional magnetic resonance imaging) scanning, shooting earmuffs were used. Participants viewed visual stimuli on a back-projection screen. The auditory stimuli were identical to those presented in a previous study<sup>37</sup> via an air-conductive tube to participants. Presentation software (ver. 10.2;

Neurobehavioral Systems) was used to generate auditory and visual stimuli on a Windows computer. In addition, the participants generated their responses using a keypad (Current Designs Inc., Philadelphia, PA, USA).

**Procedures.** The sequence of stimulus presentation is shown in Fig. 1. For each trial, a fixation cross was initially presented for 300 ms in the centre of the screen. A neutral stimulus with a straight eye gaze (gaze trail) or transverse lines (arrow trail) was subsequently presented at the location. After 350 ms, a cue stimulus (gaze or arrow) in the right or left direction was presented in the centre of the screen. The stimulus onset asynchrony (SOA) between the auditory target and cue was fixed to 200 ms. Subsequently, an auditory stimulus target (voice or tone sound) was presented in the left or right ear for 150 ms through headphones. Consistent with previous studies<sup>37–39</sup>, the participants were asked to answer quickly and precisely whether or not they heard the auditory target on the left or right side of the headphones by pressing the corresponding key on the switch keypad using their dominant index or middle fingers, respectively. Response times (RT) were measured in each trial. A standard procedure for the Posner's cueing paradigm removed cue stimuli before a target stimulus appeared on the display. However, when using a facial gaze as a cue, many studies e.g. refs 40 and 41 also implemented a modified cueing paradigm in which the cue remained on the screen until a response was obtained or a set time had elapsed. For this study, we designed a contextual processing condition between the cue and target. To establish a strong and obvious contextual relationship between the cue and target, the cue remained until a response was obtained or 1000 ms had elapsed. The targets appeared randomly on the same or opposite side of the cue direction when the cue was directed left or right. The target appeared at the cued location in 50% of the trials. The participants were told that the cue did not predict the target location and were instructed to fixate on the centre of the screen in each trial.

The fMRI analysis relied on a within-subject three factorial design, with the cue condition (gaze or arrow), context condition (i.e., the congruence of meaning between the cue and target, which could be congruent (social gaze and social voice or arrow and tone) or incongruent (social gaze and tone or arrow and social voice)), with validity condition (valid or invalid) as the repeated factors. Sixty trials were performed under each condition. Our experimental design was based on a mixed block/event-related paradigm, facilitating a more complete utilisation of the BOLD signal and enabling a deeper interpretation of how the regions of the brain function on multiple timescales<sup>42</sup>. Consistent with a previous study<sup>43</sup>, alternating blocks of experimental trials of cue condition and blocks of baseline were presented. Within the condition blocks, congruence trials were presented in a pseudorandom event-related distribution.

**MRI acquisition.** The images were acquired using a 3.0-T Trio Tim Scanner-vision whole-body MRI system (Siemens, Erlangen, Germany) to measure activation using a head coil. The functional images comprised 33 consecutive slices parallel to the anterior-posterior commissure plane, covering the entire brain. A T2\*-weighted gradient-echo planar imaging (EPI) sequence was used with the following parameters: TR = 2000 ms, TE = 30 ms, flip angle = 90°, field of view = 220 × 220 mm, matrix size = 64 × 64, and voxel size = 3.4 × 3.4 × 3.5 mm<sup>3</sup>. The slices covered most of the brain, including the entire temporal cortex, but excluding the most inferior parts of the cerebellum. We also acquired high-resolution isotropic T1-weighted images (TR = 1900 ms, TE = 2.52 ms, flip angle = 9°, field of view = 250 × 250 mm, 176 sagittal slices, voxel size = 1 × 1 × 1 mm<sup>3</sup>).

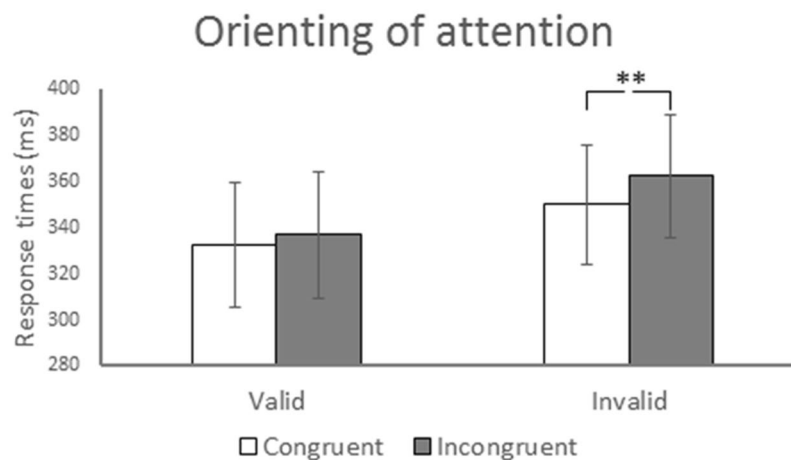
**Behavioural data analysis.** The data were analysed using the SPSS software package (ver. 21.0). Incorrect responses (1.76% of the trials) and RT of less than 100 ms or more than 1000 ms were excluded from the RT analysis (1.18% of the trials), and trials in which a response occurred prior to the target onset were also excluded. The mean RT under conditions was calculated for each participant. The mean RT was analysed using a three-way analysis of variance (ANOVA) with cue (gaze, arrow), context (congruent, incongruent), and validity (valid, invalid) as within-participant factors. To examine whether an interaction was significant, if present, a follow-up simple main effect (i.e. assessing the effect of each independent variable at each level of the other independent variable) analysis was conducted to interpret the result.

**Image data analysis.** Data preprocessing and statistical analyses were performed using the Statistical Parametric Mapping software package (SPM12; Wellcome Department of Cognitive Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk/spm/software/spm12>) implemented in MATLAB 2013b (Math Works). The functional images from each run were realigned using the first scan as a reference to correct for head movements. The movement parameters generated during spatial realignment indicated that all subjects moved less than 2 mm during the course of the trial. The T1 anatomical image was preprocessed using an intensity inhomogeneity correction. Then, T1 anatomical images were coregistered to the first scan of the functional images. Next, the coregistered T1 anatomical image was normalised to the Montreal Neurological Institute space using a unified segmentation-spatial normalisation approach<sup>44</sup>. The parameters from this normalisation process were subsequently applied to each of the functional images. Finally, these spatially normalised functional images were resampled to a voxel size of 2 × 2 × 2 and were spatially smoothed in three dimensions using an 8-mm full-width-at-half-maximum Gaussian kernel.

We used random-effects analyses<sup>45</sup> to identify significantly activated voxels exhibiting interesting effects. First, we performed a single-subject analysis<sup>46</sup>. The BOLD response was modelled as the neural activity, convolved with a canonical haemodynamic response function (HRF), which yielded regressors in a general linear model (GLM) for each condition. We used a high-pass filter comprising a discrete cosine basis function with a cut-off period of 128 to eliminate the artefactual low-frequency trend. To correct the global fluctuation related to motion artefacts, global scaling was conducted. Serial autocorrelation, assuming an AR (1) (first-order autoregressive) model, was estimated from the pooled active voxels with a restricted maximum likelihood procedure and used to whiten the data and design matrix<sup>47</sup>.

| Cue   | Validity | Context   |           |      |             |           |      |
|-------|----------|-----------|-----------|------|-------------|-----------|------|
|       |          | Congruent |           |      | Incongruent |           |      |
|       |          | <i>M</i>  | <i>SD</i> | %E   | <i>M</i>    | <i>SD</i> | %E   |
| Gaze  | Valid    | 328.3     | 28.3      | 0.15 | 330.3       | 27.9      | 0.15 |
|       | Invalid  | 344.0     | 26.7      | 0.31 | 356.2       | 26.7      | 0.34 |
| Arrow | Valid    | 337.1     | 26.3      | 0.15 | 343.4       | 27.2      | 0.12 |
|       | Invalid  | 355.9     | 25.3      | 0.27 | 368.8       | 26.7      | 0.26 |

**Table 1.** Mean response times (ms), standard deviations, and percent errors (%E) as a function of cue, context, and validity.



**Figure 2.** Response times (RT) results in attentional orienting. Mean (with SE) RT presented for valid and invalid conditions as a function of a cue type condition (gaze or arrow). \*\* $p < 0.01$ .

The contrast images from the first-level analyses from all subjects were subsequently used for the second-level group statistics. First, for each participant, the data were best fitted at every voxel using a combination of effects of interest. These data were delta functions representing the onsets of the eight conditions, given by the crossing of our  $2 \times 2 \times 2$  factorial design: cue (gaze, arrow)  $\times$  context (congruent, incongruent)  $\times$  validity (valid, invalid), convolved with the SPM12 haemodynamic response function. Second, based on the behavioural results, a  $2 \times 2 \times 2$  (cue  $\times$  context  $\times$  validity) factorial ANOVA was used to investigate the relationship between behavioural results and brain activation. Based on a methods analysis<sup>48</sup>, the statistical maps exhibited a spatial extent threshold at  $p < 0.05$ , family-wise error (FWE)-corrected for multiple comparisons, and an intensity threshold at  $p < 0.001$ , uncorrected for multiple comparisons at the whole-brain level was used to protect against false-positive activations. The peak voxels of clusters exhibiting reliable effects are reported in MNI coordinates. We had an a priori hypothesis regarding the activity of contextual processing in TPJ and STS, and the influence of contextual processing in dorsal and ventral frontoparietal networks. Based on anatomical masks using the WFU Pickatlas tool, a small-volume correction was also employed to the a priori regions of interest, attributed to the anatomical structures in left/right hemisphere of the STS with BA21, 22, the IPS and SPL with BA5, 7, the FEF with BA8, and the IFG with BA44, 45, 47, separately. Consistent with the whole-brain level, we used small-volume correction at a voxel spatial extent threshold at  $p < 0.05$ , FWE corrected, and an intensity threshold at  $p < 0.001$ , uncorrected for multiple comparisons. Finally, to quantify neural responses with the influence of attentional orienting under context conditions, we used the MarsBaR software package<sup>49</sup> to extract percentage changes in BOLD signals for congruent and incongruent contexts under valid and invalid conditions, averaged across voxels with given regions of interest (ROI) using spheres with a radius of 8 mm. Then, the means of the percent signal change (PSC) between the conditions were compared using repeated-measures ANOVA. All statistics were calculated using the SPSS software package (ver. 21).

## Results

**Behavioural results.** The data pertaining to errors did not reveal any significant main effect or interaction (all  $p > 0.05$ ), thus indicating that the participants suffered no speed-accuracy trade-off (Table 1).

The mean RT under each condition are listed in Table 1, and the mean differences in RT between the invalid and valid conditions are shown in Fig. 2. A three-factor repeated-measures ANOVA was used to analyse the RT. The analysis revealed a main effect of cue ( $F(1, 21) = 12.412$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.371$ ), with faster responses under the eye gaze (339.7 ms) versus arrow (351.3 ms) condition. In addition, we also observed a significant main effect of context ( $F(1, 21) = 8.213$ ,  $p = 0.009$ ,  $\eta_p^2 = 0.281$ ), with faster responses under congruent (341.3 ms) versus

incongruent (349.7 ms) conditions, and validity ( $F(1, 21) = 25.247, p < 0.001, \eta_p^2 = 0.546$ ), with faster responses under valid (334.8 ms) versus invalid (356.2 ms) conditions.

A significant interaction of context  $\times$  validity was observed ( $F(1, 21) = 4.907, p = 0.038, \eta_p^2 = 0.189$ ), but no significant interaction was detected for cue  $\times$  context ( $F(1, 21) = 0.166, p = 0.688, \eta_p^2 = 0.008$ ), cue  $\times$  validity ( $F(1, 21) = 0.048, p = 0.828, \eta_p^2 = 0.002$ ), or cue  $\times$  context  $\times$  validity ( $F(1, 21) = 0.108, p = 0.746, \eta_p^2 = 0.005$ ).

The *post hoc* test revealed a significant difference between context conditions under invalid conditions ( $p = 0.004$ ) but not under valid conditions ( $p = 0.182$ ) with a faster response for congruent (349.9 ms) versus incongruent (362.5 ms) under invalid conditions, indicating an RT benefit for targets that match the context (e.g. social) of the cue under invalid conditions but not under valid conditions. This result suggests that the disengagement of attention from cued locations is facilitated through contextual processing. These findings demonstrated that attentional orienting is modulated through contextual processing only under invalid conditions, enabling the investigation of the neural substrates underlying the behavioural response of the contextual processing between the cue and target in attentional orienting induced by gaze and arrow cues.

**Supplementary analysis of the influenced by the gender.** Because previous studies have reported that gaze-triggered orienting is different between genders (e.g., refs 50 and 51), we added gender (male, female) as between-participant factor to supplement the influence of the gender based on the main (3-way) ANOVA of RT data analysis, although 22 participants were recruited with gender unbalance including 9 women and 13 men. The results found a significant interaction between gender and the validity ( $F(1, 20) = 9.591, p = 0.006, \eta_p^2 = 0.324$ ) but not between gender and other factors (all  $F(1, 20) \leq 2.54, p > 0.1$ ). However, the *post hoc* test did not reveal a difference between genders under valid or invalid conditions (both  $p > 0.1$ ), although a faster response was observed for valid compared with invalid conditions in both male (323.5 vs. 335.6 ms,  $p = 0.017$ ) and female (351.1 vs. 386.0 ms,  $p < 0.001$ ) participants. Thus, the results of the present study suggest that attentional orienting through contextual processing was not influenced by the gender of the participants.

**fMRI results.** Next, based on the behavioural results, we investigated the patterns of brain activation associated with cross-modal attention. In the primary analysis, we performed 2 cue conditions (gaze, arrow)  $\times$  2 context conditions (congruent, incongruent)  $\times$  2 validity conditions (valid, invalid) repeated-measures ANOVA.

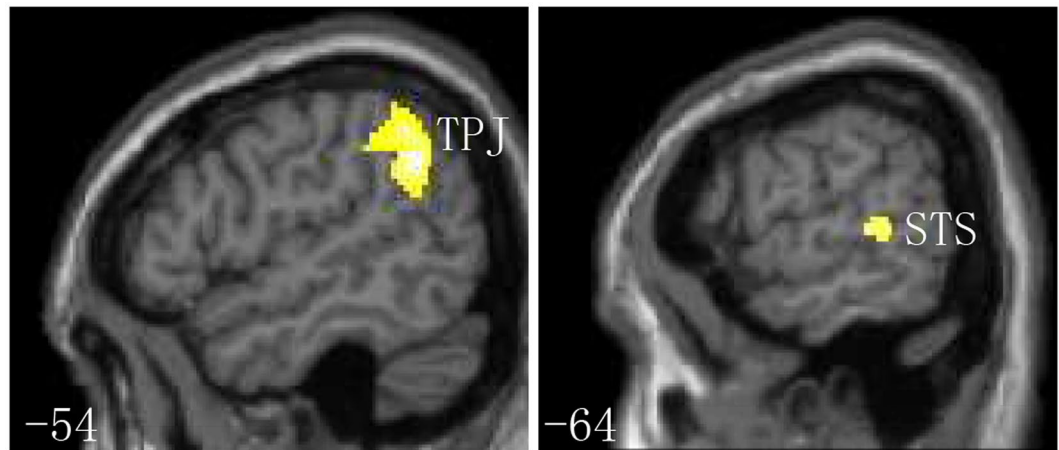
**Main effects of cue, validity, and context.** In a whole-brain analysis, the gaze trials evoked significantly greater activity than arrow trials in a many clusters of voxels. One of these clusters included the fusiform gyrus (BA 19), extending from the extra-striate visual areas into the occipital and temporal cortices (Supplementary Fig. S1, Table S1). In contrast, greater activity for arrow than for gaze trials was observed in the right hemisphere of temporal lobe, including the middle temporal gyrus (BA 37), and the left hemisphere of occipital lobe, including the middle occipital gyrus (BA 19) (Supplementary Fig. S2, Table S2). These results were consistent with previous evidence<sup>21</sup>, thus indicating that gaze versus arrow cues increased activation in various occipital and temporal areas, whereas the reverse contrast evoked activation in occipital regions.

Furthermore, in a whole-brain analysis, invalid gaze and arrow cues evoked a significantly larger response than valid gaze and arrow cues in the left frontal hemisphere and the limbic system, including the inferior and middle frontal gyrus, and the anterior cingulate (Supplementary Fig. S3, Table S3). These results were also consistent with those of a previous study<sup>16</sup> that revealed common activity in the IFG by conjunction analyses to response gaze and arrow cues in attentional orienting. However, activation was not observed by valid versus invalid gaze and arrow cues.

To highlight the neural underpinnings of RT modulated through contextual processing, we examined the difference between congruent and incongruent contextual meanings of the cue-target. The results of the whole-brain analysis indicated that the congruent condition evoked a significantly smaller response than the incongruent condition in the left hemisphere parietal, including TPJ (BA 40). Additionally, anatomical region-based small-volume corrections revealed significant activation in the temporal lobe, including STS (BA21/22). (Fig. 3, Table 2) These findings indicated that activity in left TPJ and STS regions was associated with the contextual processing of the cue-target, but this activation was not observed in the congruent versus the incongruent condition. Furthermore, to examine the differences between gazes and arrows in contextual processing, we investigated whether the neural activity in these regions differed between gaze and arrow cues. The results revealed no significant difference between gaze and arrow cues, indicating that comparable neural activity was elicited by contextual processing between gaze and arrow, thereby influencing valid and invalid orienting within attentional orienting. Next, we assessed the cue condition, focusing on its influence in contextual processing for both gaze and arrow cues in attentional orienting.

**Interaction of context and validity conditions.** A 2 (context: congruent, incongruent)  $\times$  2 (validity: valid, invalid) ANOVA was performed to investigate the influence of activation by the contextual relationship of cue-target in attentional orienting networks. The results of the whole-brain analysis revealed a significant interaction in left hemisphere TPJ (BA 40). Additionally, anatomical region-based small-volume corrections revealed significant activation in the left hemisphere IFG (BA 47) regions (Fig. 4, Table 3).

**ROI analysis.** The results of the interaction were expanded using an ROI-based analysis. Figure 4 and Table 4 present the location and pattern of the response in all ROIs in which a signal change was extracted. These responses were located in left TPJ and IFG regions. The PSC in these regions was analysed using a 2 (context: congruent, incongruent)  $\times$  2 (validity: valid, invalid) repeated-measures ANOVA. A significant interaction was observed in the left TPJ and IFG regions. Moreover, the *post hoc* test revealed that PSC was smaller when the contextual meaning of cue-target was congruent vs. incongruent in all regions under invalid conditions (all  $p < 0.05$ ) but not under valid conditions (Fig. 4, Table 4). These results indicated that the influence of contextual



**Figure 3.** In response to incongruent versus congruent context conditions, exploratory whole-brain analysis indicating that the left TPJ is significantly activated, and small-volume-correction analysis showing that the left STS is significantly activated based on an anatomical mask. A voxel-wise spatial extent threshold  $p < 0.05$ , FWE-corrected, and an intensity threshold  $p < 0.001$ , uncorrected, were used.

| Side                              | Area     | Region                   | BA    | Coordinates |     |    | Z-value | P (FWE)         | P (FWE)      | P (uncorr)   | Cluster size<br>(mm <sup>3</sup> ) |
|-----------------------------------|----------|--------------------------|-------|-------------|-----|----|---------|-----------------|--------------|--------------|------------------------------------|
|                                   |          |                          |       | x           | y   | z  |         | (cluster level) | (peak level) | (peak level) |                                    |
| Exploratory whole-brain analysis  |          |                          |       |             |     |    |         |                 |              |              |                                    |
| L                                 | Parietal | Temporoparietal junction | 40    | -54         | -50 | 34 | 4.696   | 0.001           | 0.021        | 0.000        | 819                                |
| Small-volume corrections analysis |          |                          |       |             |     |    |         |                 |              |              |                                    |
| L                                 | Temporal | Superior temporal sulcus | 21/22 | -64         | -34 | 6  | 3.943   | 0.028           | 0.005        | 0.000        | 41                                 |

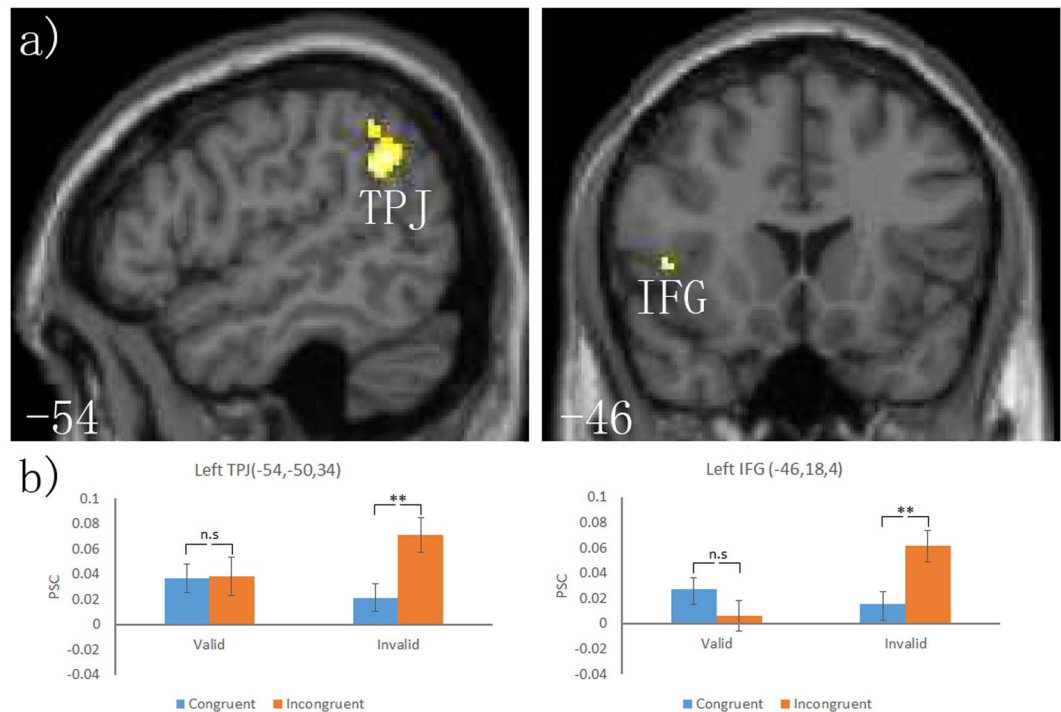
**Table 2.** Main effect of incongruent condition: incongruent > congruent. BA = Brodmann's area; FWE = family-wise error; a voxel-wise spatial extent threshold at  $p < 0.05$ , FWE corrected, and an intensity threshold at  $p < 0.001$ , uncorrected.

processing on the neural activity for attentional orienting was observed under invalid conditions but not under valid conditions.

## Discussion

We examined attentional orienting by gaze and arrow cues under a localising task in which participants were asked to indicate whether the target (voice and tone) was heard on the left or right side of the headphones. The combination of cues and targets varied across trials, reflecting contextual relationship processing, where the contextual meaning of the cue-target was congruent (social gaze and social voice and arrow and tone) or incongruent (social gaze and tone and arrow and social voice). Although the behavioural results showed no difference between gaze and arrow cues, a RT benefit was observed for targets matching the context of the cue under the invalid condition. This finding suggested that a disengagement of attention from cued locations was facilitated through contextual processing. Previous studies<sup>23,25</sup> have demonstrated that attentional orienting can be influenced by contextual processing when targets were presented in a specific context (e.g., colour or emotion) for gaze or arrow cues. Compared with a previous study<sup>23</sup> in which attentional orienting based on a contextual effect for gaze and arrow was investigated in the context of colour (i.e., schematic white/black eyes as the cue and a black square as the target), the present study examined attentional orienting through contextual processing using facial gaze and voice, which seems to more closely resemble a real-world environment. Furthermore, although another study<sup>25</sup> examined attentional orienting through gaze cues influenced by emotional context for 80 different images (e.g., a chimney image) as targets, the present study only manipulated two sounds as targets, potentially easing the establishment of a pairing between the cue and target. That is, the pairing of gaze and voice was easily established, and arrow and tone represented the other pair. Thus, the present study observed attentional orienting through contextual processing when using gaze and arrow as cues. Based on these findings, the results of the present study extended those of previous studies<sup>23,25</sup>, indicating that attentional orienting through centrally presented cues, irrespective of cue characteristics (e.g. social or non-social), could also be modulated by contextual relationship processing between the cue and target.

Importantly, consistent with the behavioural results, the main analyses of fMRI data revealed that neural substrates were not different in response to the contextual relationship processing of cue-target for gaze and arrow cues. Previous studies<sup>28-31</sup> have revealed that neural substrates in the regions of the TPJ and STS were associated with contextual processing. Consistently, to highlight the neural underpinnings of RT modulated by contextual processing, the results of the present study also demonstrated that the left STS and bilateral TPJ were specifically



**Figure 4.** (a) In response to the interaction between context and validity conditions, exploratory whole-brain analysis showing left TPJ significantly activated, and small-volume-correction analysis showing left IFG significantly activated based on an anatomical mask. A voxel-wise spatial extent threshold  $p < 0.05$ , FWE corrected, and an intensity threshold  $p < 0.001$ , uncorrected, were used. (b) Mean ( $\pm$ SE) and percent signal changes (PSC) in the left hemisphere TPJ and IFG regions are shown. These areas are overlaid on the mean normalised structural MRI from all subjects in this study. n.s:  $p > 0.05$ ; \*\* $p < 0.01$ .

| Side                              | Area     | Region                   | BA | Coordinates |     |    | Z-value | P (FWE)         | P (FWE)      | P (uncorr)   | Cluster size |
|-----------------------------------|----------|--------------------------|----|-------------|-----|----|---------|-----------------|--------------|--------------|--------------|
|                                   |          |                          |    | x           | y   | z  |         | (cluster level) | (peak level) | (peak level) |              |
| Exploratory whole-brain analysis  |          |                          |    |             |     |    |         |                 |              |              |              |
| L                                 | Parietal | Temporoparietal junction | 40 | -54         | -50 | 34 | 3.962   | 0.01            | 0.345        | 0.000        | 391          |
| Small-volume corrections analysis |          |                          |    |             |     |    |         |                 |              |              |              |
| L                                 | Frontal  | Inferior frontal gyrus   | 47 | -46         | 18  | 4  | 3.356   | 0.04            | 0.025        | 0.000        | 5            |

**Table 3.** Interaction between context and validity conditions. BA = Brodmann's area; FWE = family-wise error; a voxel-wise spatial extent threshold at  $p < 0.05$ , FWE corrected, and an intensity threshold at  $p < 0.001$ , uncorrected.

| Regions defined by locating local maxima |                 |                               |                                 |
|--|-----------------|-------------------------------|---------------------------------|
| Region                                   | Interaction $F$ | Valid incongruent > congruent | Invalid incongruent > congruent |
| Left TPJ                                 | 4.392*          | 0.004                         | 15.308**                        |
| Left IFG                                 | 10.082**        | 3.715                         | 8.956**                         |

**Table 4.** ROI results. ROIs represent previously examined areas that exhibited a significant interaction between context and validity conditions in a  $2 \times 2$  ANOVA (with a voxel-wise spatial extent threshold at  $p < 0.05$ , FWE corrected, and an intensity threshold at  $p < 0.001$ , uncorrected). \* $p < 0.05$ , \*\* $p < 0.01$ .

involved in contextual processing under both gaze and arrow conditions. The results also indicated that these regions were weakly activated when the relationship of cue-target was congruent (the expected target matching the context of the cue) compared with incongruent (the expected target non-matching the context of the cue). We suggest that these regions may be the locations of an inhibitory mechanism, which enhanced neural activity to suppress the processing of incongruent predictions for targets from cue stimuli. Given that the activity of these regions did not differ between social gaze and arrow cues, we further suggest that a comparable neural system was elicited by contextual processing for gaze and arrow cues and further influenced the attentional process. This idea



is consistent with previous studies<sup>12–18</sup> that demonstrated the differences in attention to social and non-social cues were quantitative rather than qualitative.

Interaction analyses in behavioural results indicated that a different pattern of attentional orienting by contextual processing was elicited for valid and invalid conditions. That is, a RT benefit was observed for targets matching the context (e.g. social) of the cue under invalid conditions, but not under valid conditions. We propose that the different patterns between valid and invalid conditions may be influenced through the overlap of the time window between contextual processing and attentional orienting. Electrophysiological studies have demonstrated that the activity of temporal staging differed between valid and invalid conditions<sup>52–54</sup>. These studies reported an amplitude enhanced at P1 (a positive component at occipital electrode sites between 70 and 100 ms post-target onset) in attentional orienting with gaze as the cue under valid versus invalid conditions, whereas a greater amplitude at P3 (a positive component at central/parietal/midline electrode sites between 300 and 500 ms post-target onset) was observed under invalid versus valid conditions. However, in previous studies<sup>55, 56</sup>, the N300 (a negative component at frontal electrode sites at approximately 300 ms) - N400 (a negative component at central/parietal electrode sites at approximately 400 ms) wave reflected an updating of context information. For example, Demiral *et al.*<sup>55</sup> observed a stronger N300-N400 effect elicited through a semantic context when the contextual scene was presented before the target in a spatial attention task. Given that the pattern of attentional orienting by contextual processing differed between valid and invalid conditions, we suggest that this finding may be influenced through associations with the activities of temporal staging between the influence of attentional orienting (i.e. the mechanism of valid and invalid conditions) and the contextual processing component. Compared with the early stage (70–100 ms) under valid conditions, the time window of processing overlapped with that of contextual information and attentional orienting under invalid conditions at a later stage (300–500 ms), in which these processes could be integrated to suppress a violation of expectancies (the expected target matching the context of the cue) when the contextual relationship of cue-target was incongruent. In addition, we speculated that varying SOA conditions may influence expectations, and the patterns of attentional orienting by contextual processing for valid and invalid conditions could be modulated in the present experiment. That is, if the target was presented at a short SOA prior to the expectation of the subject, then attentional orienting by contextual processing may not be influenced under valid or invalid conditions, whereas if the target was presented at a long SOA after the expectation of the subject, then attentional orienting by contextual processing might be influenced under invalid conditions.

Furthermore, the results of the interaction analyses in fMRI also demonstrated the influence of contextual processing on neural activity for attentional orienting under invalid but not valid conditions. Such neural activities were observed in the left hemisphere TPJ and IFG, an area in the ventral frontoparietal network that may be responsible for invalidity orienting (for reviews, see refs 10 and 11). As mentioned above, analyses of the fMRI data revealed that the left STS and TPJ were involved in the contextual processing of the relationship between the cue and target. The activity in the left TPJ region overlapped in processing the contextual relationship of cue-target and invalidity orienting in attention. Additionally, consistent with the pattern of neural activity for contextual processing in the left STS and TPJ, we observed that when the relationship of cue-target was congruent versus incongruent based on ROI analyses, less activity was observed in all of these brain regions. This result suggests that from the left TPJ to the IFG in the ventral frontoparietal network, neural signals for contextual processing were transferred to invalidity orienting in attention, in which the intrinsic connection pathway is present among these regions<sup>57</sup>. Compared with incongruent conditions, we suggest that the lower activity in the ventral frontoparietal network when the contextual processing of the relationship between cue and target is congruent under invalid conditions may reflect a disengagement of attention from cued locations at a lower cost, which is readily elicited. This idea may explain the behavioural data obtained under invalid conditions, indicating that a lower cost is associated with processing when the contextual processing of the relationship between the cue and target was congruent versus incongruent; thus, participants could disengage attention from the cued location to rapidly capture a target.

**Implications of the present study.** Previous behavioural studies<sup>23–26</sup> had demonstrated that attentional orienting by gaze or arrow cues could be modulated through contextual processing. Consistent with these studies, the behavioural results in the present study also revealed that RT in attentional orienting was modulated through contextual processing. In particular, we observed that attentional orienting was modulated through contextual processing under invalid conditions. Given that the present study identified the influence of the neural substrates by contextual processing under invalid conditions in ventral frontoparietal networks, we suggest that this finding may account for the behavioural data regarding attentional orienting through contextual processing based on the neurocognitive architecture.

In addition, a behavioural study demonstrated impaired attentional orienting when the cue-target relationship is weak (i.e., incongruent context) in individuals with autism spectrum disorder (ASD)<sup>27</sup>, a finding that raises a question regarding whether individuals with ASD exhibit impairment of gaze-triggered attention because activity is impaired in the neural mechanism in the ventral frontoparietal network. Given that impaired gaze-triggered attention may impede and differentially affect the development of the ability to understand the mental state of another individual in social communication<sup>1</sup>, we suggest that an atypical function in the ventral frontoparietal network, particularly in the processing of contextual information, may be associated with the atypical development of social cognition, further suggesting an important direction for future studies combining brain imaging and treatment interventions for social processing deficits in individuals with ASD.

**Limitations.** First, we tested attentional orienting by contextual processing using gaze and arrow as cues with two types of targets (social voice and tone) under only visual-auditory cross-modal conditions. Given the complexity of real life, future studies should examine attentional orienting by contextual processing using two types

of targets under visual-visual unimodal or visual-tactile cross-modal conditions in which attentional orienting by contextual processing may also play an important role.

Second, the present study involved two types of contextually related cues and targets. In contrast with faces and voices, which are immediately paired in a congruent context, the pairing of the tone and arrow might be influenced through the increasing number of trials throughout the experiment. To evaluate this possibility, all experimental blocks were divided into two parts (first and last half of the block), and a four-factor repeated-measures ANOVA (block  $\times$  cue  $\times$  context  $\times$  validity) was used to analyse the RTs. Given that a significant 4-way interaction was observed ( $F(1, 21) = 5.09, p = 0.04, \eta_p^2 = 0.195$ ), two 3-way repeated-measures ANOVA (block  $\times$  cue  $\times$  context) was performed under valid and invalid conditions separately. Although no significant interaction was detected for block conditions under invalid conditions, (all  $p > 0.1$ ), we observed that the main effect of context ( $F(1, 21) = 11.74, p = 0.003, \eta_p^2 = 0.36$ ) was significant with faster response to congruent than incongruent conditions (349.0 vs. 361.5 ms), thus indicating that contextual processing between the cue and target was immediately established at both gaze and arrow pairings. However, under valid conditions, although we observed no significant main effect of context ( $F(1, 21) = 1.41, p = 0.25, \eta_p^2 = 0.063$ ), a significant interaction of block  $\times$  cue  $\times$  context was observed ( $F(1, 21) = 12.67, p = 0.002, \eta_p^2 = 0.38$ ). The *post hoc* test revealed a significant difference between context conditions ( $p = 0.045$ ) with a faster response to congruent than incongruent conditions (327.4 vs. 346.9 ms) when using arrows as cues in the last half of the block, thereby indicating that attentional orienting by contextual processing could be influenced by increasing the number of trials throughout the experiment under valid conditions when the tone target was matched with arrow cue. In the present study, attentional orienting was modulated through contextual processing under invalid conditions. Thus, we suggest that this effect may be elicited through immediately established pairing between arrow and tone, rather than increasing the number of trials in the experiment. However, future studies should investigate the mechanism of how contextual processing is influenced by increasing the number of trials through the experiment.

Finally, in the present study, we directly contrasted the context conditions by gaze and arrow cues in invalid and valid trials to reveal differences in attentional orienting. Previous studies<sup>15, 20, 21</sup> manipulated a neutral cue (e.g. direct gaze) as a baseline condition to examine differences in the neural mechanisms between valid and invalid attentional orienting conditions. However, compared with a non-directional arrow as a neutral cue, Engell *et al.*<sup>19</sup> suggested that a direct gaze, as a neutral cue, was perceived as directional rather than non-directional, which was problematic in terms of comparing social versus non-social cueing in an fMRI study. Future research should investigate the need for a baseline condition in which the neural gaze and arrow cues involve no spatial information and have the same effect on neural activity, such as closed eyes and non-directional arrows.

## Conclusions

In this study, we observed that the response time in attentional orienting by gaze and arrow cues was modulated through contextual processing between the cue and target when contextually congruent and incongruent under invalid conditions in behavioural studies. Additionally, on the neural level, activity in the left TPJ and STS was observed with attentional orienting by gaze and arrow cues in response to contextual processing of the relationship between the cue and target. However, we did not observe any difference in the neural substrates between social gaze and arrows by contextual processing in attentional orienting. This finding adds further evidence in support of the notion that the differences in attention to social and non-social cues are quantitative rather than qualitative. Importantly, both behavioural and fMRI results indicated that the influence of contextual processing on neural activity for attentional orienting occurred under invalid conditions. Such an increase was observed in the ventral frontoparietal network when the cue and target were incongruent rather than congruent. This finding may provide an explanation for the behavioural data regarding attentional orienting by contextual processing based on the neurocognitive architecture.

## References

1. Baron-Cohen, S. *Mindblindness: An Essay on Autism and Theory of Mind*. (MIT press, 1995).
2. Posner, M. Orienting of attention. *Q. J. Exp. Psychol.* **32**, 3–25 (1980).
3. Friesen, C. K. & Kingstone, A. The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychon. Bull. Rev.* **5**, 490–495 (1998).
4. Bayliss, A. P. & Tipper, S. P. Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *Br. J. Psychol.* **96**, 95–114 (2005).
5. Marotta, A., Casagrande, M. & Lupiáñez, J. Object-based attentional effects in response to eye-gaze and arrow cues. *Acta Psychol.* **143**, 317–321 (2013).
6. Stevens, S. A., West, G. L., Al-Aidroos, N., Weger, U. W. & Pratt, J. Testing whether gaze cues and arrow cues produce reflexive or volitional shifts of attention. *Psychon. Bull. Rev.* **15**, 1148–1153 (2008).
7. Tipples, J. Orienting to counterpredictive gaze and arrow cues. *Percept. Psychophys.* **70**, 77–87 (2008).
8. Yan, T. *et al.* Target object moderation of attentional orienting by gazes or arrows. *Atten. Percept. Psychophys.* **78**, 2373–2382 (2016).
9. Sato, W., Kochiyama, T., Uono, S. & Yoshikawa, S. Automatic attentional shifts by gaze, gestures, and symbols. *Psychologia.* **53**, 27–35 (2010).
10. Corbetta, M., Patel, G. & Shulman, G. L. The reorienting system of the human brain: From environment to theory of mind. *Neuron* **58**, 306–324 (2008).
11. Corbetta, M. & Shulman, G. L. Spatial neglect and attention networks. *Annu. Rev. Neurosci.* **34**, 569–599 (2011).
12. Brignani, D., Guzzon, D., Marzi, C. A. & Miniussi, C. Attentional orienting induced by arrows and eye-gaze compared with an endogenous cue. *Neuropsychologia.* **47**, 370–381 (2009).
13. Callejas, A., Shulman, G. L. & Corbetta, M. Dorsal and ventral attention systems underlie social and symbolic cueing. *J. Cogn. Neurosci.* **26**, 63–80 (2014).
14. Greene, D., Mooshagian, E., Kaplan, J., Zaidel, E. & Iacoboni, M. The neural correlates of social attention: automatic orienting to social and nonsocial cues. *Psychol. Res.* **73**, 499–511 (2009).

15. Joseph, R. M., Fricker, Z. & Keehn, B. Activation of frontoparietal attention networks by non-predictive gaze and arrow cues. *Soc. Cogn. Affect Neurosci.* **10**, 294–301 (2014).
16. Sato, W., Kochiyama, T., Uono, S. & Yoshikawa, S. Commonalities in the neural mechanisms underlying automatic attentional shifts by gaze, gestures, and symbols. *NeuroImage* **45**, 984–992 (2009).
17. Tipper, C. M., Handy, T. C., Giesbrecht, B. & Kingstone, A. Brain responses to biological relevance. *J. Cogn. Neurosci.* **20**, 879–891 (2008).
18. Tipples, J., Johnston, P. & Mayes, A. Electrophysiological responses to violations of expectation from eye gaze and arrow cues. *Soc. Cogn. Affect Neurosci.* **8**, 509–514 (2013).
19. Engell, A. D. *et al.* Differential activation of frontoparietal attention networks by social and symbolic spatial cues. *Soc. Cogn. Affect Neurosci.* **5**, 432–440 (2010).
20. Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R. & Hämäläinen, H. Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage* **33**, 406–413 (2006).
21. Lockhofen, D. E. L., Gruppe, H., Ruprecht, C., Gallhofer, B. & Sammer, G. Hemodynamic response pattern of spatial cueing is different for social and symbolic cues. *Front. Hum. Neurosci.* **8**, 912, doi:10.3389/fnhum.2014.00912 (2014).
22. Birmingham, E. & Kingstone, A. Human social attention. *Prog. Brain Res. Elsevier: Volume 176* (ed Srinivasan, Narayanan) 309–320 (Elsevier, 2009).
23. Ristic, J., Wright, A. & Kingstone, A. Attentional control and reflexive orienting to gaze and arrow cues. *Psychon. Bull. Rev.* **14**, 964–969 (2007).
24. Pratt, J. & Hommel, B. Symbolic control of visual attention: The role of working memory and attentional control settings. *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 835–845 (2003).
25. Bayliss, A. P., Schuch, S. & Tipper, S. P. Gaze cueing elicited by emotional faces is influenced by affective context. *Vis. Cogn.* **18**, 1214–1232 (2010).
26. Kuhn, G. & Tipples, J. Increased gaze following for fearful faces. It depends on what you're looking for! *Psychon. Bull. Rev.* **18**, 89–95 (2011).
27. Zhao, S., Uono, S., Yoshimura, S., Kubota, Y. & Toichi, M. Can gaze-cueing be helpful for detecting sound in autism spectrum disorder? *Res. Autism Spectr. Disord.* **7**, 1250–1256 (2013).
28. Geng, J. J. & Vossel, S. Re-evaluating the role of TPJ in attentional control: Contextual updating? *Neurosci. Biobehav. Rev.* **37**, 2608–2620 (2013).
29. Weidner, R., Krummenacher, J., Reimann, B., Müller, H. J. & Fink, G. R. Sources of top-down control in visual search. *J. Cogn. Neurosci.* **21**, 2100–2113 (2008).
30. Noppeney, U., Josephs, O., Hocking, J., Price, C. J. & Friston, K. J. The effect of prior visual information on recognition of speech and sounds. *Cereb. Cortex.* **18**, 598–609 (2008).
31. Sass, K., Sachs, O., Krach, S. & Kircher, T. Taxonomic and thematic categories: Neural correlates of categorization in an auditory-to-visual priming task using fMRI. *Brain Res.* **1270**, 78–87 (2009).
32. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia.* **9**, 97–113 (1971).
33. Said, C. P., Sebe, N. & Todorov, A. Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion.* **9**, 260–264 (2009).
34. Zebrowitz, L. A., Kikuchi, M. & Fellous, J. M. Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *J. Pers. Soc. Psychol.* **98**, 175–189 (2010).
35. Montepare, J. M. & Dobish, H. The contribution of emotion perceptions and their overgeneralizations to trait impressions. *J. Nonverbal Behav.* **27**, 237–254 (2013).
36. Uono, S. & Hietanen, J. K. Eye contact perception in the west and east: A cross-cultural study. *PLoS One* **10**, e0118094, doi:10.1371/journal.pone.0118094 (2015).
37. Li, C., Chen, K., Han, H., Chui, D. & Wu, J. An fMRI study of the neural systems involved in visually cued auditory top-down spatial and temporal attention. *PLoS One* **7**, e49948, doi:10.1371/journal.pone.0049948 (2012).
38. Zhao, S., Uono, S., Yoshimura, S. & Toichi, M. Is impaired joint attention present in non-clinical individuals with high autistic traits? *Mol. Autism.* **6**, 67, doi:10.1186/s13229-015-0059-3 (2015a).
39. Zhao, S., Uono, S., Yoshimura, S. & Toichi, M. Self make-up: The influence of self-referential processing on attention orienting. *Sci. Rep.* **5**, 14169, doi:10.1038/srep14169 (2015b).
40. Ristic, J., Friesen, C. K. & Kingstone, A. Are eyes special? It depends on how you look at it. *Psychon. Bull. Rev.* **9**, 507–513 (2002).
41. Ristic, J. *et al.* Eyes are special but not for everyone: The case of autism. *Brain Res. Cogn. Brain Res.* **24**, 715–718 (2005).
42. Petersen, S. E. & Dubis, J. W. The mixed block/event-related design. *NeuroImage* **62**, 1177–1184 (2012).
43. Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N. & Dale, A. M. Stochastic designs in event-related fMRI. *NeuroImage* **10**, 607–619 (1999).
44. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).
45. Holmes, A. P. & Friston, K. J. Generalisability, random effects and population inference. *NeuroImage* **7**, S754 (1998).
46. Friston, K. J. *et al.* Analysis of fMRI time-series revisited. *NeuroImage* **2**, 45–53 (1995).
47. Friston, K. J. *et al.* Classical and bayesian inference in neuroimaging: Applications. *NeuroImage* **16**, 484–512 (2002).
48. Woo, C. W., Krishnan, A. & Wager, T. D. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage* **91**, 412–419 (2014).
49. Brett, M., Anton, J., Valabregue, R. & Poline, J. Region of interest analysis using an SPM toolbox. *Presented at the 8<sup>th</sup> International Conference on Functiona Mapping of the Human Brain*. Available on CD-ROM in *NeuroImage* **16** (2002).
50. Bayliss, A. P., di Pellegrino, G. & Tipper, S. P. Sex differences in eye gaze and symbolic cueing of attention. *Q. J. Exp. Psychol.* **58**, 631–650 (2005).
51. Ciardo, F., Ricciardelli, P., Lugli, L., Rubichi, S. & Iani, C. Eyes keep watch over you! Competition enhances joint attention in females. *Acta Psychol.* **160**, 170–177 (2015).
52. Schuller, A.-M. & Rossion, B. Spatial attention triggered by eye gaze increases and speeds up early visual activity. *NeuroReport.* **12**, 2381–2386 (2001).
53. Schuller, A.-M. & Rossion, B. Perception of static eye gaze direction facilitates subsequent early visual processing. *Clin. Neurophysiol.* **115**, 1161–1168 (2004).
54. Schuller, A.-M. & Rossion, B. Spatial attention triggered by eye gaze enhances and speeds up visual processing in upper and lower visual fields beyond early striate visual processing. *Clin. Neurophysiol.* **116**, 2565–2576 (2005).
55. Demiral, S. B., Malcolm, G. L. & Henderson, J. M. ERP correlates of spatially incongruent object identification during scene viewing: Contextual expectancy versus simultaneous processing. *Neuropsychologia* **50**, 1271–1285 (2012).
56. Hamm, J. P., Johnson, B. W. & Kirk, I. J. Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clin. Neurophysiol.* **113**, 1339–1350 (2002).
57. Vossel, S., Geng, J. J. & Fink, G. R. Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *Neuroscientist.* **20**, 150–159 (2013).

## Acknowledgements

The authors would like to thank the participants who were involved in the study. This study was financially supported by a Grant-in-Aid for Overseas JSPS Fellows (1515307), and the National Natural Science Foundation of China (grant numbers 61473043, 61305142, 81671776 and 11302139), Beijing Municipal Science & Technology Commission (grant numbers Z161100002616020) and The Capital Medical University Fundamental and Clinical Foundations of China (grant numbers 16JL-L08).

## Author Contributions

S.Z. and C. L. collected the data; S.Z. analysed data, prepared the figures and drafted the manuscript. S. U. made a substantial contribution to writing and revising the manuscript; All authors participated in designed the experiments, contributed to the manuscript development, read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-03104-1](https://doi.org/10.1038/s41598-017-03104-1)

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017