VISUAL COGNITION
https://doi.org/10.1080/13506285.2025.2462043

# A perceptual advantage for social groups in interactive configurations

Clara Colombatto [a], Francesca Capozzi [b], Victoria Fratino[c] and Jelena Ristic [d]

[a]Department of Psychology, University of Waterloo, Waterloo, ON, Canada; [b]Department of Psychology, University of Quebec a Montreal (UQAM), Montreal, QC, Canada; [c]Department of Psychology, Concordia University, Montreal, QC, Canada; [d]Department of Psychology, McGill University, Montreal, QC, Canada

**ABSTRACT**
Humans have a long-standing evolutionary history of group belonging. Our visual system should thus be tuned to detect social groups, especially those in interactive or "core configurations," where group members face each other. Past work shows that two individuals are detected more efficiently when they are facing toward (vs. away from) each other. Here we tested whether this facing advantage extends to small social groups of three, or triads. In three preregistered experiments, participants searched for a facing group (among non-facing ones) or a non-facing group (among facing ones). Facing groups were found faster than non-facing ones, demonstrating a perceptual advantage for groups in core configurations (Experiment 1). This advantage persisted in inverted displays, suggesting a role for cues to body orientation (Experiments 2 and 3). Human perception is thus well-tuned to detect not just prototypical dyadic interactions, but interactive configurations more generally, facilitating efficient processing of complex social information.

**ARTICLE HISTORY**
Received 25 July 2024
Accepted 28 January 2025

## Introduction

The human cognitive and perceptual systems have evolved in environmental conditions favouring social interactions and cooperation (Adolphs, 2009). Indeed, past work has shown that perception is well-equipped for detecting the presence of humans (Allison et al., 2000; Capozzi & Ristic, 2018). For example, when viewing natural scenes, we more readily notice human bodies (New et al., 2007) and faces (Ro et al., 2001) relative to other objects. This preference for social information emerges early in development (Farroni et al., 2002) and has been associated with brain networks specialized for processing a variety of social cues (Kanwisher et al., 1997; Pitcher & Ungerleider, 2021).

While work in social perception has explored how humans detect and attend to people when they are presented in isolation, in everyday life we typically encounter multiple people at the same time: much of what we perceive are not single individuals, but rather *groups* of individuals. Although the size of social groups varies from intimate circles to large crowds, humans typically congregate in small groups of three to five people (Dunbar et al., 1995). These small social groups are characterized by a "core configuration" with individuals facing each other – an arrangement that promotes group development, differentiation of social roles, and group coordination (Caporael, 1997). Accordingly, while investigations of social perception have so far largely focused on studying individuals, more recent views tend to shift the focus from "person perception" to "people perception" (Alt & Phillips, 2022), with groups understood as social units that transcend individual members (Levine, 2018; Shamay-Tsoory et al., 2019).

Given the pervasiveness and importance of social groups, here we hypothesized that human perception may be attuned to efficiently detect groups in core configurations (Phillips et al., 2014). That is, when looking at scenes containing groups of individuals, the visual system may prioritize cues to their configurations (i.e., face-to-face body orientation), resulting in more efficient detection of social groups. Past work provides some support for this hypothesis. First, a long-standing tradition in perception has shown that when viewing sets of objects, we rapidly and automatically extract their summary properties (see Whitney & Yamanashi Leib, 2018). This capacity for ensemble representation also applies to people, as in when we perceive the mean gazing direction (Sweeny & Whitney, 2014)

**CONTACT** Clara Colombatto clara.colombatto@uwaterloo.ca; Jelena Ristic jelena.ristic@mcgill.ca

or emotional expression (Haberman & Whitney, 2007) of a crowd of faces. Summary statistics of crowds such as sex ratio have even been found to guide perceivers' social attitudes such as perceived threat (Alt et al., 2019) and sense of belonging (Goodale et al., 2018).

Beyond statistical properties, when seeing people we also readily perceive the interactions between them. For example, hands belonging to separate individuals are attended to as a single object when they are performing a social handshake but not when performing a similar non-social action (e.g., a handshake with inverted hands; Shen et al., 2016; Yin et al., 2018). Further, two people are recognized more accurately (Papeo et al., 2017), found faster in search displays (Papeo et al., 2019), and remembered more accurately (Vestner et al., 2019) when the individuals are facing toward each other (a "facing advantage"; for a review, see Papeo, 2020). The perception of small groups (e.g., of three to five people), however, may follow different principles relative to groups of two due to the increased quantity of social information (Ristic & Capozzi, 2022) and additional degrees of freedom in the possible interactions among individual members (Lehmann-Willenbrock et al., 2017). These quantitative and qualitative differences between dyads and larger social groups raise a fundamental question about the nature of search processes in social contexts, namely whether these computations are based on a prototypical representation of dyadic interactions, or a more general computation of interactive possibilities amongst individuals.

In the current work, we thus asked whether small social groups of three would be prioritized in perceptual processing, especially when the members are arranged in core group configurations. In Experiment 1, we assessed search performance for small social groups in core configurations (i.e., with group members facing each other; facing triads) compared to non-core configurations (i.e., with group members facing away from each other; non-facing triads). In Experiments 2 and 3, we investigated possible mechanisms underlying social group detection by examining search performance for groups in core and non-core configurations situated within inverted displays.

## Experiment 1

Experiment 1 examined whether facing groups of three are detected more efficiently when the individuals are positioned in a core configuration (i.e., facing each other) compared to a non-core configuration (i.e., facing away from each other). Participants viewed displays containing four or eight triads. On half the blocks, they searched for a facing triad within non-facing triads, and on the remaining half they searched for a non-facing triad within facing triads. Following our hypothesis of a perceptual advantage for social groups, we expected to find faster response times for facing relative to non-facing target triads.

### Methods

The preregistered methods and analyses are available at https://osf.io/k5x64. Anonymized raw data and analysis scripts for this and the subsequent experiments are available at https://osf.io/mz7d2.

### Participants

Participants ($N = 139$) were recruited through McGill University's participant pool and Prolific Academic (prolific.-com) in exchange for course credit and monetary compensation, respectively. Participants were excluded according to the preregistered plan if they had a mean overall response accuracy lower than 65% ($N = 17$), or a total percentage of valid trials lower than 75% of all trials ($N = 17$). The final sample was 105 participants ($N_{women} = 78$, $N_{men} = 23$, $N_{other} = 4$, Mean age = 25.91). All reported normal or corrected to normal vision. The sample size was based on a power analysis of pilot data conducted in R using the pwr package (v1.3.0), which revealed that a sample size of 105 would be sufficient to achieve 90% power to detect the effect size of interest in a paired-samples, two-tailed *t*-test with an alpha level of 0.05, with the effect size of interest being the within-subjects difference in response times for locating facing vs. non-facing triads ($d_z = 0.32$).

Participants completed the experiment online on their own devices, via a website link controlled by custom software written in HTML, JavaScript, CSS, and PHP, using the jsPsych library (De Leeuw, 2015). Their browser window was automatically put in full-screen mode at the beginning of the experiment, and participants were required to remain in full-screen mode for the entire experiment.

### Stimuli

Figure 1(a) shows examples of facing and non-facing triads. The stimuli were created using Daz3D studio

(Daz Productions, Salt Lake City, UT). Six different facing triads were created by slightly modifying individual poses within each triad (Figure 1(a), left panel; see also Papeo et al., 2019). For each facing triad, a corresponding non-facing triad was created by rotating each individual in place by 180° (Figure 1(a), right panel). Each triad was rendered in greyscale from a constant viewpoint. The size of each triad was determined separately for each participant such that the width of the stimulus would be 8% of the width of their browser window (when in full-screen mode), and the height was adjusted proportionally.

To ensure that the relative positioning of individuals within triads was equated across different versions of all facing and non-facing triads, we confirmed that these stimuli did not differ in (a) the horizontal distance between extremities of neighbouring individuals (Wilcoxon Signed-Rank Test $W = 5$, $z = 1.15$, $p > 0.05$); and (b) the horizontal distance between the centre of neighbouring individuals ($W = 6$, $z = 0.94$, $p > 0.05$).

Sample search displays are depicted in Figure 1(b). Each triad could be placed in one of 16 possible spatial locations on the screen, located on the perimeters of two imaginary ellipses centred on the screen – an inner one with its radius width and height spanning 25% of participants' screen width and height, respectively, and an outer one with its radius width and height spanning 40% of participants' screen width and height, respectively. Each triad could be located at 30°, 60°, 120°, 150°, 210°, 240°, 300°, and 330° along each ellipsis' radius, with two constraints: (1) following Papeo et al. (2019), the target triad was always positioned along the perimeter of the inner ellipsis, such that it would be detected more easily given the higher spatial resolution of foveal vision; and (2) there was always an equal number of triads in each screen quadrant (i.e., one per quadrant in displays with set size four and two per quadrant in displays with set size eight). The poses of each triad were sampled randomly from the six possible poses without replacement for arrays of set size four, and with two poses repeated for arrays of set size eight.

### Procedure

Each trial began with a central black fixation cross (72 × 72 px) shown on a white background for 500 ms. Next, the search array was displayed and remained visible until response, or until 2500 ms had elapsed. Participants were instructed to indicate quickly and accurately whether the target triad was located on the left or right side of the screen by pressing "b" or "h" on their keyboard. Response type-key assignment was randomized across participants. The search array was followed by a blank screen for 500 ms, after which the next trial began.

### Design

Target type was manipulated across blocks. On half of the blocks, participants searched for a facing triad in an array of non-facing triads; on the remaining half of blocks, they searched for a non-facing triad in an array of facing triads (Target Type: Facing or Non-Facing). Blocks with Facing and Non-Facing targets were interleaved, and the order of blocks was randomized for each participant. Instructions indicating the response target were displayed at the beginning of each block. The remaining factors were intermixed and randomly manipulated within blocks. The search array could contain three or seven distractors in addition to the target (Set Size: 4 or 8), and the targets could appear in one of eight possible positions along the perimeter of the inner ellipsis (Target Position: 30°, 60°, 120°, 150°, 210°, 240°, 300°, or 330°).

This within-subjects design resulted in 2 block conditions (corresponding to 2 Target Types) and 16 trial conditions (corresponding to 2 Set Sizes × 8 Target Positions). Each of these 16 unique trial combinations was repeated twice in a block (in a randomized order), and each participant completed eight blocks (in a randomized order) for a total of 256 trials per participant. At the end of each block, participants could take a short break. The experiment was preceded by eight practice trials divided into two blocks, one with facing targets and one with non-facing targets.

### Analyses

Following the preregistered plan, we removed trials with missed responses (6.33% of each participant's trials, on average), inaccurate responses (8.86% of each participant's trials, on average; note that these trials were not excluded for accuracy analyses), and response times above or below 2.5 standard deviations from each participant's mean (1.57% of each participant's remaining trials, on average). There was no speed-accuracy trade-off ($r(103) = 0.06$, $p = 0.512$).

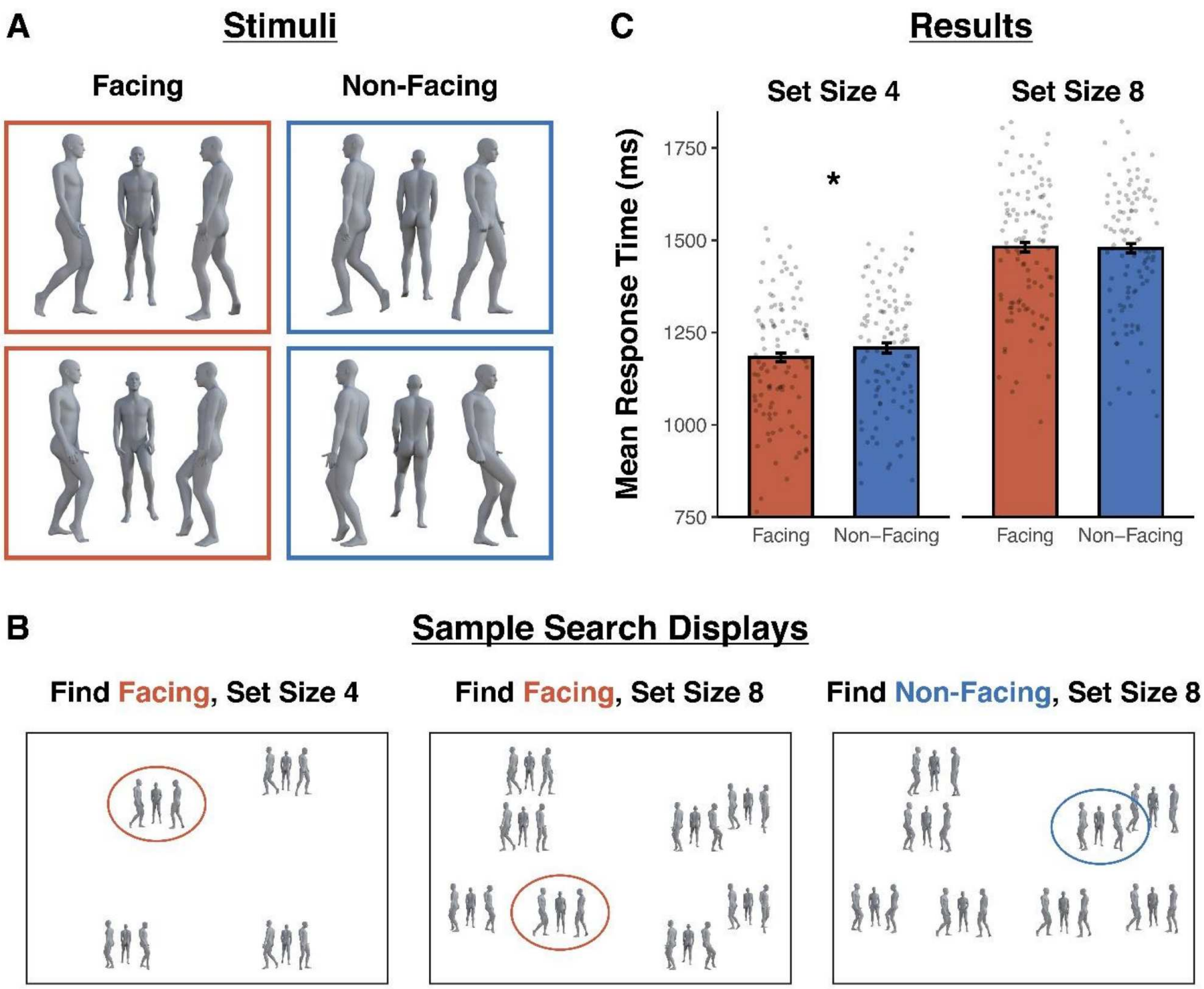


**Figure 1.** Stimuli and results from Experiment 1. (a) Example facing and non-facing triads, from two of six different poses used in the experiment. (b) Sample search displays for facing targets among non-facing distractors, and non-facing targets among facing distractors; note that stimuli have been enlarged and targets highlighted for visualization purposes. (c) Average response times for facing and non-facing triads. Dots represent participant means, error bars reflect 95% confidence intervals, subtracting out the shared variance. *$p < 0.05$.

Accuracy and mean response times on correct trials were examined in two separate 2 × 2 repeated measures ANOVAs with factors Target Type (Facing or Non-Facing) and Set Size (4 or 8).

### *Results*

Overall accuracy was high ($M = 92.10\%$). The accuracy analyses revealed a main effect of Set Size ($F(1, 104) = 48.35$, $p < 0.001$, $\eta_p^2 = 0.32$), with higher accuracy for the smaller set size (mean [$M$] = 93.50%, standard error [SE] = 0.39) compared to the larger set size ($M = 90.69\%$, SE = 0.53). No other effects reached significance (all $F$s < 1.44, all $p$s > 0.233).

Response time data are illustrated in Figure 1(c). There was a main effect of Set Size ($F(1, 104) = 1504.42$, $p < 0.001$, $\eta_p^2 = 0.94$), with faster response times for the smaller set size ($M = 1195.28$ ms, SE = 15.06) compared to the larger set size ($M = 1479.79$ ms, SE = 16.86). While there was no reliable main effect of Target Type ($F(1, 104) = 1.40$, $p = 0.239$, $\eta_p^2 = 0.01$), there was a reliable interaction between Set Size and Target Type ($F(1, 104) = 8.50$, $p = 0.004$, $\eta_p^2 = 0.08$). Post-hoc tests corrected for multiple comparisons using the Bonferroni method confirmed that response times were significantly faster for Facing ($M = 1182.39$ ms, SE = 15.86) than for Non-Facing targets ($M = 1208.16$ ms, SE = 16.15;

$t(104) = 2.38$, $p = 0.019$, $d_z = 0.23$) at the smaller set size, but not at the larger set size (Facing $M =$ 1481.32 ms, SE = 17.47; Non-Facing $M = 1478.26$ ms, SE = 17.92; $t(104) = -0.28$, $p = 0.776$, $d_z = -0.03$).

### Discussion

Experiment 1 revealed that participants were faster to find facing triads (i.e., triads in their core configuration) among non-facing triads compared to non-facing triads among facing triads. This search advantage was larger for smaller search displays, which is consistent with the notion that individuation of social information becomes more difficult as the number of individuals increases (Ristic & Capozzi, 2022). These results thus suggest that triads in their core configurations are prioritized in visual search, reflecting the importance of detecting social groups in our environment.

## Experiment 2

The search advantage for facing triads in their core configuration may be supported by several mechanisms. One possibility is that our visual system responds preferentially to individuals in core configurations because of their familiarity: while some non-facing configurations are relatively familiar (e.g., multiple people all facing forward in a classroom or a theatre), it is relatively uncommon to encounter individuals who are facing exactly away from each other (Colombatto et al., 2020). Another possibility is that perception responds to social cues more flexibly, and cues to group configurations (e.g., whether the individuals are facing toward vs. away from each other) can be extracted even when the groups are presented in more unfamiliar ways. To test these possibilities, in Experiment 2 we asked participants to search for facing and non-facing triads, but with all search displays inverted – a manipulation known to disrupt the perception of complex social properties but to preserve the perception of cues such as body orientation (Tanaka et al., 2022; Vestner et al., 2022). If the search advantage for facing groups depended on perceiving the groups in familiar upright configurations, we expected that it would disappear with stimulus inversion. Alternatively, a search advantage for facing inverted triads would suggest that visual cues to core configurations such as body orientation might also influence the perception of social groups.

### Methods

The preregistered methods and analyses can be viewed at https://osf.io/t9xwg.

Participants (total $N = 189$) were recruited according to the same procedure as in Experiment 1. Participants were excluded according to the preregistered plan if they had a mean overall accuracy lower than 65% ($N =$ 52), or a total percentage of valid trials lower than 75% ($N = 32$). The final sample was 105 participants ($N_{women} = 75$, $N_{men} = 26$, $N_{other} = 4$, Mean age = 24.49). All participants had normal or corrected-to-normal vision. This sample size was chosen to match Experiment 1.

This experiment was identical to Experiment 1, except that all stimuli were inverted (i.e., rotated 180° around the horizontal axis), as shown in Figure 2(a,b).

Following the preregistered plan, we removed trials with missed responses (8.09% of each participant's trials, on average), inaccurate responses (10.69% of each participant's trials, on average; note that these trials were not excluded for accuracy analyses), and response times above or below 2.5 standard deviations from each participant's mean (1.07% of each participant's remaining trials, on average). There was no speed-accuracy trade-off ($r(103) = -0.01$, $p = 0.945$). As in Experiment 1, accuracy and mean response times on correct trials were examined in two separate $2 \times 2$ repeated measures ANOVAs with factors Target Type (Facing or Non-Facing) and Set Size (4 or 8).

### Results

The accuracy analyses revealed a main effect of Set Size ($F(1, 104) = 73.40$, $p < 0.001$, $\eta_p^2 = 0.41$), with higher accuracy for the smaller set size ($M = 92.09\%$, SE = 0.45) compared to the larger set size ($M =$ 88.14%, SE = 0.55). No other effects reached significance (all $Fs < 2.15$, $ps > 0.145$). Once again, overall accuracy was high ($M = 90.12\%$).

Response time data are illustrated in Figure 2(c). There was a main effect of Set Size ($F(1, 104) =$ 1681.19, $p < 0.001$, $\eta_p^2 = 0.94$), with faster response times for the smaller set size ($M = 1274.77$ ms, SE = 14.32) compared to the larger set size ($M =$ 1582.17 ms, SE = 14.13). In contrast to Experiment 1, there was no interaction between Set Size and

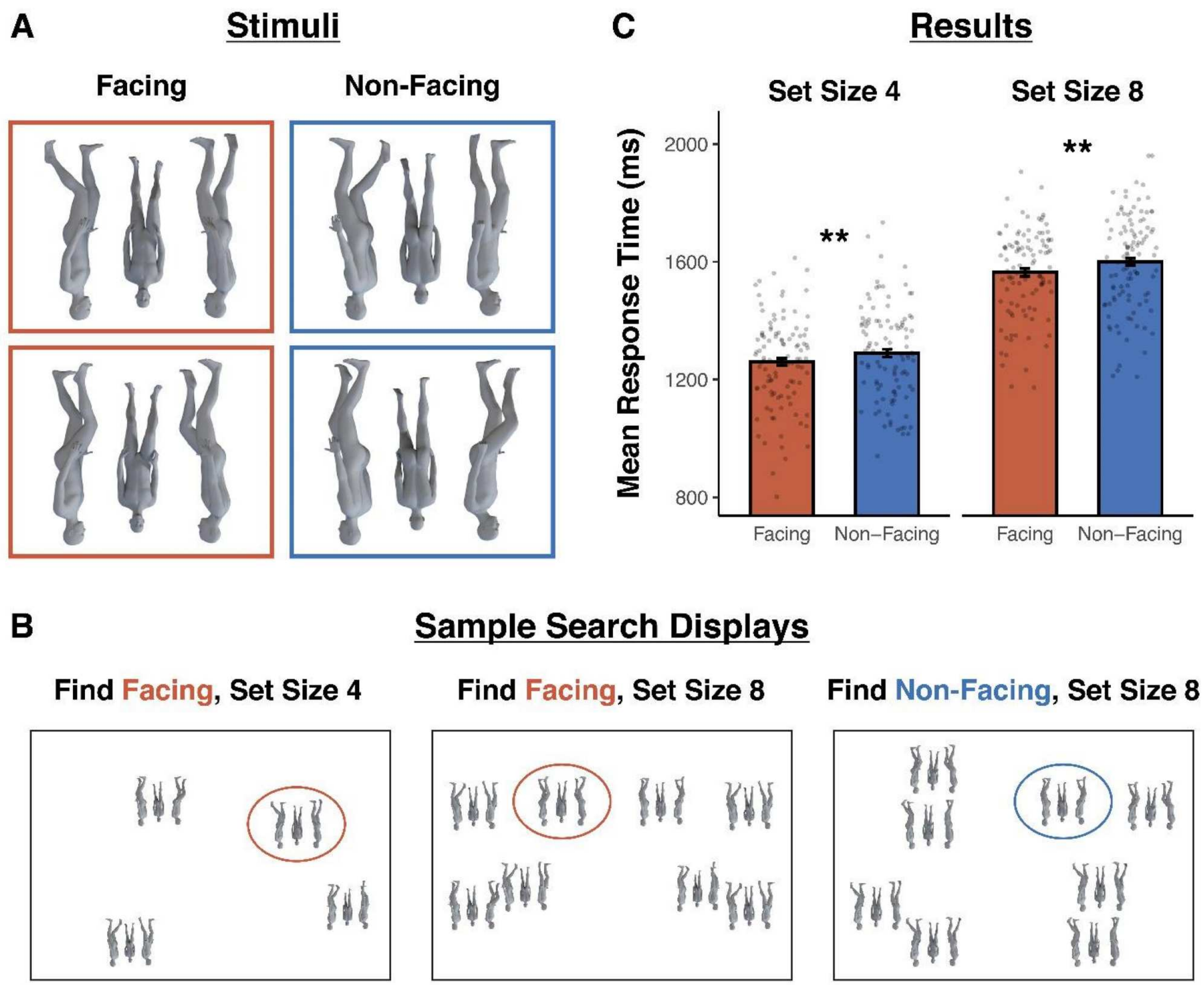


**Figure 2.** Stimuli and results from Experiment 2. (a) Example facing and non-facing inverted triads, from two of six different poses used in the experiment. (b) Sample search displays for facing targets among non-facing distractors, and non-facing targets among facing distractors; note that stimuli have been enlarged and targets highlighted for visualization purposes. (c) Average response times for facing and non-facing inverted triads. Dots represent participant means, error bars reflect 95% confidence intervals, subtracting out the shared variance. $^{**}p < 0.005$.

Target Type ($F(1, 104) = 0.37$, $p = 0.543$, $\eta_p^2 < 0.01$), but there was a reliable main effect of Target Type ($F(1, 104) = 12.15$, $p = 0.001$, $\eta_p^2 = 0.10$), with faster response times for Facing ($M = 1412.06$ ms, SE = 13.96) compared to Non-Facing targets ($M = 1444.88$ ms, SE = 15.04).

### *Discussion*

The results of Experiment 2 revealed that the search advantage for facing triads persisted when the search displays were inverted. A comparison of Experiments 1 and 2 (Appendix) further indicated that participants were overall slower in searching through inverted relative to upright displays, which is consistent with past work on impaired perception of inverted bodies (Reed et al., 2003; Vestner, Gray, et al., 2021). The finding that facing triads were still prioritized even in inverted search displays dovetails with the idea that the perception of social groups may not be orientation-specific but may arise from the perception of cues indicative of core configurations (e.g., facing direction), which can be perceived even in unfamiliar viewing conditions.

## Experiment 3

The results of Experiment 2 indicated that facing triads may be prioritized even in inverted displays, suggesting that the effects of core group configurations may be preserved in unfamiliar viewing conditions. However, in Experiment 2, the search arrays were displayed for a relatively long duration (up to 2500 ms), which may have reduced the disruptive effects of inversion (e.g., by allowing for mental rotation). To confirm that this did not influence the results, we re-ran the same experiment using shorter exposure durations.

### *Methods*

The preregistered methods and analyses are available at https://osf.io/74mzg. This experiment was identical to Experiment 2, except that the search arrays were displayed for half of the original duration, 1250 ms rather than 2500 ms. Given the increased task difficulty, we modified the exclusion criteria (which now also matched past similar work, e.g., Papeo et al., 2019), to exclude: (1) participants whose mean accuracy and/or mean response times on correct trials were above or below 2.5 standard deviations from the group mean; and (2) for response times analyses, trials with incorrect or missed responses, and trials with response times above or below 2 standard deviations from the individual's mean.

Following these criteria, from the initial sample of $N = 108$, we excluded 1 participant for low accuracy and 2 participants for low response times. The final sample was 105 participants ($N_{women} = 90$, $N_{men} = 14$, $N_{other} = 1$, Mean age = 20.38). All participants had normal or corrected-to-normal vision. This sample size was chosen to match Experiments 1 and 2.

Following the preregistered plan, we removed trials with missed responses (10.24% of each participant's trials, on average), inaccurate responses (22.06% of each participant's trials, on average), and response times above or below 2 standard deviations from each participant's mean (3.63% of each participant's remaining trials, on average). As the data indicated a speed-accuracy trade-off ($r(103) = 0.63$, $p < 0.001$), following our preregistered plan, we conducted our main analyses on inverse efficiency scores calculated for each participant and condition (i.e., individual mean response time divided by proportion correct for each condition). These scores were examined in a 2 × 2 repeated measures ANOVA with factors Target Type (Facing or Non-Facing) and Set Size (4 or 8).

### *Results*

The inverse efficiency analyses revealed a main effect of Set Size ($F(1, 104) = 200.07$, $p < 0.001$, $\eta_p^2 = 0.66$), with faster responses in the smaller set size ($M =$ 1416.40 ms, SE = 23.36) compared to the larger set size ($M = 1718.88$ ms, SE = 30.67). More importantly, there was a reliable main effect of Target Type ($F(1, 104) = 6.35$, $p = 0.013$, $\eta_p^2 = 0.06$), with faster response times for Facing ($M = 1538.52$ ms, SE = 28.54) compared to Non-Facing targets ($M = 1596.76$ ms, SE = 26.65). There was no interaction between Set Size and Target Type ($F(1, 104) = 0.15$, $p = 0.702$, $\eta_p^2 < 0.01$). As detailed in the Appendix, these patterns were also reflected in the separate analyses of response times and accuracy.

### *Discussion*

The results from Experiment 3 replicated Experiment 2, as the search advantage for facing triads emerged again in inverted displays – but now also when the displays were viewed for a short duration, which limited the use of strategies that may balance the disruptive effects of inversion (e.g., mental rotation). The increased difficulty of this experiment also eliminated the ceiling performance in accuracy from Experiments 1 and 2, such that the effects of core configurations now emerged in both accuracy and response time, consistent with past work (Papeo et al., 2019; Vestner et al., 2019; Vestner, Over, et al. 2021). An exploratory comparison of Experiments 2 and 3 (Appendix) confirmed that the effect of target type was reliable and stable across both experiments. The advantage for facing triads is thus relatively unaffected by disruptions in group configural processing and may be influenced by visual cues that remain discernible even in unfamiliar viewing conditions.

## General discussion

While the study of social perception is typically focused on how we perceive other individuals, in everyday life we often encounter groups of multiple people. Indeed, humans preferentially engage with and congregate in groups comprised of about 3–5 people (Dunbar et al., 1995), and these small social groups have specific evolutionary advantages as suggested by work in anthropology (Caporael, 1997). Social groups typically exhibit an organized structure and specific boundaries – resulting in a "core configuration" characterized by the individuals of the group facing each other. Given the pervasiveness and evolutionary importance of core social group configurations, we hypothesized that perception might be tuned to detect social groups in such arrangements. To address this question, we measured participants' performance as they searched for facing groups among non-facing ones, and for non-facing groups among facing ones.

Overall, facing groups were found faster than non-facing groups, suggesting a perceptual advantage for social groups in their core configuration. In Experiment 1, this search advantage was stronger for smaller search displays, when three distractors were present, relative to larger search displays, when seven distractors were present. This suggests that participants can more effectively engage specialized or automatic mechanisms for processing socially relevant stimuli in conditions of low attentional demands. On the other hand, larger set sizes resulted in longer response times; this suggests that participants searched through the displays serially, rejecting distractors in turn, which resulted in longer response times as the number of distractors increased. The reduction of the facing advantage at this larger set size suggests that search advantages for social groups may be reduced under high attentional demands and in serial search conditions. Future work may further explore how different attentional demands and search strategies may result in differential engagement of social processes. Further, the absence of a search advantage in larger search displays suggests that the search advantage for facing groups is likely not driven by lower-level stimulus properties, which are preserved (if not more evident) in larger and more crowded displays. More importantly, the interaction of the facing advantage with set size also suggests that the perception of social cues may decrease in crowded scenes with a larger number of people, such that individuation of social groups may become more difficult with crowding (Ristic & Capozzi, 2022; Yan et al., 2024).

Interestingly, the search advantage for facing groups persisted even when the search displays were inverted in Experiments 2 and 3. This may be surprising given that stimulus inversion is typically thought to disrupt configural processing, resulting in impaired recognition of bodies including their postures (Reed et al., 2006) and other social properties (e.g., Pavlova & Sokolov, 2000). The current data are overall consistent with this general recognition impairment in inverted displays, as target detection was reliably slowed down for inverted displays in Experiment 2 compared to the upright displays in Experiment 1. However, facing triads were still prioritized even in inverted orientations (in Experiments 2 and 3), and despite shortened presentation times (in Experiment 3). This result is contrary to prior findings with dyads, where the facing advantage instead appeared to be disrupted in inverted displays (Papeo et al., 2017; Vestner, Over, et al., 2021). This discrepancy may be due to the increased power of the current experiments, given the larger sample sizes compared to past studies. However, these differential results in inverted displays may also reflect a greater role for configural processing in dyads compared to triads. For example, the added complexity of larger social groups may facilitate processing of spatial relations even in inverted displays, or of cues indicating body direction such as fingers (Ariga & Watanabe, 2009) or feet (Dalmaso, 2023) which can be perceived even when presented in unfamiliar orientations. This dovetails with studies of face perception, which suggest that inversion may disrupt face processing quantitatively rather than qualitatively (Sekuler et al., 2004).

An important consideration in the interpretation of the current results is the similarity between the triadic stimuli employed here and traditional dyadic stimuli used in prior research. While the stimuli employed in the current experiments were designed to represent groups of three individuals, through spatial alignment and implied engagement, these triads may also be perceived as featuring a dyad along with an additional individual. Therefore, it remains unclear whether the facing advantage for social groups arises from representations of all individuals, or from the representation of dyadic relationships which persists even when embedded within larger groups. It is worth noting however that contrary to dyads (Papeo et al., 2017; Vestner et al., 2019), in the present studies the advantage for facing triads remained robust even in inverted displays. This points to a more pronounced role for configural processing in the perception of dyads compared to triads, and suggests that the perception of triadic stimuli may involve distinct processes from the perception of dyadic stimuli. This possibility is also consistent with recent investigations of larger groups (up to eight individuals), where the facing advantage diminished with increasing group size, despite the presence of dyadic relationships (Yan et al., 2024). Future work could further investigate the role of relationships within larger groups by exploring alternative group configurations (e.g., with two individuals facing each other and a third facing away), and further elucidating the specificity of triadic processing in visual cognition.

The current experiments focused on a specific kind of perceptual prioritization, testing the speed at which different groups were located in search arrays. These findings raise the possibility that other perceptual processes may also be sensitive to social configuration. For example, the search advantage we have observed here may reflect a sort of "perceptual grouping" (Beck, 1966; Treisman, 1982; Wagemans et al., 2012) wherein separate individuals are perceived as a single perceptual unit when arranged in configurations that suggest a social group (e.g., when they face each other). Future work may thus explore whether signatures of perceptual grouping discovered in the non-social domain, such as impairments in the individuation of group members, may also apply to this kind of social grouping.

The current study is based on a comparison of search times for facing targets (among non-facing distractors) and non-facing targets (among facing distractors). The search advantages for facing triads may thus reflect either faster localization of facing targets or faster rejection of non-facing distractors. The relative contribution of target detection vs. distractor rejection may be further explored by using a third group configuration that facing and non-facing groups can be compared to, such as all individuals facing forward. Indeed, social groups may be arranged in a variety of configurations beyond the ones we investigated in the current experiments. For example, the back-to-back groups employed here may still be seen as having a common intention (e.g., moving away from the centre), and we would expect that perceptual advantages may also disappear in other kinds of non-core configurations, e.g., if the orientation of individuals appeared to be random. The perception of social groups may be determined by multiple other factors beyond configuration, including physical features (e.g., size or distance between members) as well as socially relevant features (e.g., demographic characteristics or interactions between members; Zhou et al., 2019). Further, the prioritization of core configurations may extend beyond social stimuli to non-social ones as even simple shapes can be arranged in ways such that they appear to be "facing" each other, and may thus invoke analogous perceptual advantages (e.g., Vestner, Over, et al., 2021; see also Colombatto, van Buren, & Scholl, 2020). Future work should thus explore the role of such properties in the perception of social groups, as well as their priority of perceptual influence and generality across social and non-social stimuli (Lockwood et al., 2020).

Finally, this work paves the way for a study of the perception of social groups beyond individuals or dyads and raises important questions about potential mechanisms driving the perception of social groups of different sizes. One possibility is that the perception of groups may involve the same processes as when processing individuals or dyads, although with an increased quantity of available social information in larger groups (Ristic & Capozzi, 2022). On the other hand, social groups may differ from smaller units not just quantitatively, but also qualitatively. While we can easily track one or two individuals, tracking individual and group dynamics in larger groups is based on both individual and aggregate group signals, and can be more effortful (Capozzi et al., 2016, 2019). At the same time, small social groups remain different from larger crowds, where individuation is no longer possible (Ristic & Capozzi, 2022). Future work may thus explore how the variations in group numerosity and social complexity affect social perception and its mechanisms, from individuals to crowds (see also Yan et al., 2024).

Overall, the current experiments demonstrate that social groups in core configurations are prioritized in visual search, suggesting that human vision is fine-tuned to detect not just individuals, but also social groups. As such, these results add to a growing literature on the social sensitivity of vision, wherein foundational aspects of human perception are influenced by visually subtle yet highly socially informative factors.

## Acknowledgments

The authors thank Salvador Soto-Faraco and members of the McGill Laboratory for Attention and Social Cognition for helpful discussions.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

The authors were supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grants 2018-04986 to JR, 2024-04579 to CC), Social Sciences and Humanities Research Council of Canada (SSHRC Insight Grant 435-2018-0361 to FC and JR), and McGill University (G. W. Stairs Memorial Fund 260766 to JR).

## Authors' contributions

CC, FC, and JR designed the research; VF implemented the stimuli; CC performed research and analyzed data; CC and JR wrote the manuscript, with feedback from all authors.

## Ethics approval

All reported studies were approved by the Research Ethics Board committee of McGill University. The procedures used in these studies adhere to the tenets of the Declaration of Helsinki.

## Consent to participate

Informed consent was obtained from all individual participants prior to participation.

## Open practices statement

Preregistrations can be viewed at https://osf.io/k5x64 (Experiment 1), https://osf.io/t9xwg (Experiment 2), and https://osf.io/74mzg (Experiment 3). Raw data and analysis code for each experiment can be viewed at https://osf.io/mz7d2.

## ORCID

*Clara Colombatto* http://orcid.org/0000-0003-3293-8741
*Francesca Capozzi* http://orcid.org/0000-0002-0062-437X
*Jelena Ristic* http://orcid.org/0000-0002-6046-9682

## References

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*(1), 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278. https://doi.org/10.1016/S1364-6613(00)01501-1

Alt, N. P., Goodale, B., Lick, D. J., & Johnson, K. L. (2019). Threat in the company of men: Ensemble perception and threat evaluations of groups varying in sex ratio. *Social Psychological and Personality Science*, *10*(2), 152–159. https://doi.org/10.1177/1948550617731498

Alt, N. P., & Phillips, L. T. (2022). Person perception, meet people perception: Exploring the social vision of groups. *Perspectives on Psychological Science*, *17*(3), 768–787. https://doi.org/10.1177/17456916211017858

Ariga, A., & Watanabe, K. (2009). What is special about the index finger?: The index finger advantage in manipulating reflexive attentional shift. *Japanese Psychological Research*, *51*(4), 258–265. https://doi.org/10.1111/j.1468-5884.2009.00408.x

Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, *1*(5), 300–302. https://doi.org/10.3758/BF03207395

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, *1*(4), 276–298. https://doi.org/10.1207/s15327957pspr0104_1

Capozzi, F., Becchio, C., Willemse, C., & Bayliss, A. P. (2016). Followers are not followed: Observed group interactions modulate subsequent social attention. *Journal of Experimental Psychology: General*, *145*(5), 531–535. https://doi.org/10.1037/xge0000167

Capozzi, F., Beyan, C., Pierro, A., Koul, A., Murino, V., Livi, S., Bayliss, A. P., Ristic, J., & Becchio, C. (2019). Tracking the leader: Gaze behavior in group interactions. *iScience*, *16*, 242–249. https://doi.org/10.1016/j.isci.2019.05.035

Capozzi, F., & Ristic, J. (2018). How attention gates social interactions. *Annals of the New York Academy of Sciences*, *1426*(1), 179–198. https://doi.org/10.1111/nyas.13854

Colombatto, C., van Buren, B., & Scholl, B. J. (2020). Gazing without eyes: A "stare-in-the-crowd" effect induced by simple geometric shapes. *Perception*, *49*(7), 782–792. https://doi.org/10.1177/0301006620934320

Dalmaso, M. (2023). Foot cues can elicit covert orienting of attention. *Psychological Research*, *87*, 2440–2448. https://doi.org/10.1007/s00426-023-01827-7

De Leeuw, J. R. (2015). Jspsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dunbar, R. I., Duncan, N. D., & Nettle, D. (1995). Size and structure of freely forming conversational groups. *Human Nature*, *6*(1), 67–78. https://doi.org/10.1007/BF02734136

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Goodale, B. M., Alt, N. P., Lick, D. J., & Johnson, K. L. (2018). Groups at a glance: Perceivers infer social belonging in a group based on perceptual summaries of sex ratio. *Journal of Experimental Psychology: General*, *147*(11), 1660–1676. https://doi.org/10.1037/xge0000450

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751–R753. https://doi.org/10.1016/j.cub.2007.06.039

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Lehmann-Willenbrock, N., Hung, H., & Keyton, J. (2017). New frontiers in analyzing dynamic group interactions: Bridging social and computer science. *Small Group Research*, *48*(5), 519–531. https://doi.org/10.1177/1046496417718941

Levine, J. M. (2018). Socially-shared cognition and consensus in small groups. *Current Opinion in Psychology*, *23*, 52–56. https://doi.org/10.1016/j.copsyc.2017.12.003

Lockwood, P. L., Apps, M. A., & Chang, S. W. (2020). Is there a 'social' brain? Implementations and algorithms. *Trends in Cognitive Sciences*, *24*(10), 802–813. https://doi.org/10.1016/j.tics.2020.06.011

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*, *104*(42), 16598–16603. https://doi.org/10.1073/pnas.0703913104

Papeo, L. (2020). Twos in human visual perception. *Cortex*, *132*, 473–478. https://doi.org/10.1016/j.cortex.2020.06.005

Papeo, L., Goupil, N., & Soto-Faraco, S. (2019). Visual search for people among people. *Psychological Science*, *30*(10), 1483–1496. https://doi.org/10.1177/0956797619867295

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. *Psychological Science*, *28*(3), 369–379. https://doi.org/10.1177/0956797616685769

Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*(5), 889–899. https://doi.org/10.3758/BF03212075

Phillips, L. T., Weisbuch, M., & Ambady, N. (2014). People perception: Social vision of groups and consequences for organizing and interacting. *Research in Organizational Behavior*, *34*, 101–127. https://doi.org/10.1016/j.riob.2014.10.001

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. *Trends in Cognitive Sciences*, *25*(2), 100–110. https://doi.org/10.1016/j.tics.2020.11.006

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, *14*(4), 302–308. https://doi.org/10.1111/1467-9280.14431

Reed, C. L., Stone, V. E., Grubb, J. D., & McGoldrick, J. E. (2006). Turning configural processing upside down: Part and whole body postures. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 73–87. https://doi.org/10.1037/0096-1523.32.1.73

Ristic, J., & Capozzi, F. (2022). Mechanisms for individual, group-based and crowd-based attention to social information. *Nature Reviews Psychology*, *1*(12), 721–732. https://doi.org/10.1038/s44159-022-00118-z

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science*, *12*(1), 94–99. https://doi.org/10.1111/1467-9280.00317

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*(5), 391–396. https://doi.org/10.1016/j.cub.2004.02.028

Shamay-Tsoory, S. G., Saporta, N., Marton-Alper, I. Z., & Gvirts, H. Z. (2019). Herding brains: A core neural mechanism for social alignment. *Trends in Cognitive Sciences*, *23*(3), 174–186. https://doi.org/10.1016/j.tics.2019.01.002

Shen, M., Yin, J., Ding, X., Shui, R., & Zhou, J. (2016). Deployment of attention on handshakes. *Frontiers in Psychology*, *7*, 681. https://doi.org/10.3389/fpsyg.2016.00681.

Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science*, *25*(10), 1903–1913. https://doi.org/10.1177/0956797614544510

Tanaka, Y., Ishikawa, K., Oyama, T., & Okubo, M. (2022). Face inversion does not affect the reversed congruency effect of gaze. *Psychonomic Bulletin & Review*, *30*(3), 974–982. https://doi.org/10.3758/s13423-022-02208-8

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194–214. https://doi.org/10.1037/0096-1523.8.2.194

Vestner, T., Gray, K. L., & Cook, R. (2021). Visual search for facing and non-facing people: The effect of actor inversion. *Cognition*, *208*, 104550. https://doi.org/10.1016/j.cognition.2020.104550

Vestner, T., Gray, K. L., & Cook, R. (2022). Sensitivity to orientation is not unique to social attention cueing. *Scientific Reports*, *12*(1), 5059. https://doi.org/10.1038/s41598-022-09011-4

Vestner, T., Over, H., Gray, K. L., Tipper, S. P., & Cook, R. (2021). Searching for people: Non-facing distractor pairs hinder the visual search of social scenes more than facing distractor pairs. *Cognition*, *214*, 104737. https://doi.org/10.1016/j.cognition.2021.104737

Vestner, T., Tipper, S. P., Hartley, T., Over, H., & Rueschemeyer, S. A. (2019). Bound together: Social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. *Journal of Experimental Psychology: General*, *148*(7), 1251–1268. https://doi.org/10.1037/xge0000545

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*(1), 105–129. https://doi.org/10.1146/annurev-psych-010416-044232

Yan, L., Colombatto, C., & Ristic, J. (2024). Search advantages for facing social groups reflect optimal interactive group sizes. *iScience*, *27*(11), 111105. https://doi.org/10.1016/j.isci.2024.111105

Yin, J., Xu, H., Duan, J., & Shen, M. (2018). Object-based attention on social units: Visual selection of hands performing a social interaction. *Psychological Science*, *29*(7), 1040–1048. https://doi.org/10.1177/0956797617749636

Zhou, C., Han, M., Liang, Q., Hu, Y. F., & Kuai, S. G. (2019). A social interaction field model accurately identifies static and dynamic social groupings. *Nature Human Behaviour*, *3*(8), 847–855. https://doi.org/10.1038/s41562-019-0618-2

## Appendix. Additional analyses

### Experiment 1 (upright) vs. Experiment 2 (inverted)

As per our preregistered plan, we examined the effect of stimulus inversion directly in a 2 × 2 × 2 mixed ANOVA with a between-subjects factor of Stimulus Orientation (Upright from Experiment 1; Inverted from Experiment 2), and within-subjects factors of Target Type (Facing or Non-Facing) and Set Size (4 or 8). First, this analysis revealed a main effect of Stimulus Orientation ($F(1, 208) = 19.21$, $p < 0.001$, $\eta_p^2 = 0.08$), with faster response times for Upright ($M = 1337.53$ ms, SE = 14.67) compared to Inverted targets ($M = 1428.47$ ms, SE = 14.67). There was also a main effect of Set Size ($F(1, 208) = 3184.70$, $p < 0.001$, $\eta_p^2 = 0.94$), with faster response times for the smaller ($M = 1235.02$ ms, SE = 10.39) compared to the larger set size ($M = 1530.98$ ms, SE = 11.00). This main effect was qualified by a two-way interaction with Stimulus Orientation ($F(1, 208) = 4.76$, $p = 0.030$, $\eta_p^2 = 0.02$), as the difference between set sizes was greater for Inverted Stimuli in Experiment 2 ($t(208) = 41.45$, $p < 0.001$, $d_z = 4.04$) compared to Upright Stimuli in Experiment 1 ($t(208) = 38.36$, $p < 0.001$, $d_z = 3.74$).

Most importantly, there was a reliable main effect of Target Type ($F(1, 208) = 10.82$, $p = 0.001$, $\eta_p^2 = 0.05$), with overall faster response times for Facing ($M = 1371.96$ ms, SE = 10.67) compared to Non-Facing Targets ($M = 1394.05$ ms, SE = 11.13). This main effect was similar in the two experiments (no interaction between Stimulus Orientation and Target Type, $F(1, 208) = 2.56$, $p = 0.111$, $\eta_p^2 = 0.01$), but was qualified by a three-way interaction with Stimulus Orientation and Set Size ($F(1, 208) = 6.36$, $p = 0.012$, $\eta_p^2 = 0.03$) which indicated that while for Inverted Stimuli in Experiment 2 there was a significant difference between Facing and Non-Facing targets for both Set Size 4 ($t(208) = 2.83$, $p = 0.005$, $d_z = 0.19$) and Set Size 8 ($t(208) = 3.33$, $p = 0.001$, $d_z = 0.23$), for Upright Stimuli in Experiment 1 the difference was significant only for Set Size 4 ($t(208) = 2.43$, $p = 0.016$, $d_z = 0.17$), and not for Set Size 8 ($t(208) = -0.28$, $p = 0.776$, $d_z = -0.02$).

### Experiment 3 accuracy and response times

There were no interactions between Set Size and Target Type (response times: $F(1, 104) = 1.23$, $p = 0.270$, $\eta_p^2 = 0.01$; accuracy: $F(1, 104) = 0.05$, $p = 0.831$, $\eta_p^2 < 0.01$). There were main effects of Set Size for both response times ($F(1, 104) = 163.17$, $p < 0.001$, $\eta_p^2 = 0.61$) and accuracy ($F(1, 104) = 108.50$, $p < 0.001$, $\eta_p^2 = 0.51$), with faster and more accurate responses in the smaller set size ($M = 1101.13$ ms, SE = 19.86; $M = 78.94\%$, SE = 1.31) compared to the larger set size ($M = 1228.99$ ms, SE = 24.73; $M = 72.56\%$, SE = 1.25). The main effect of Target Type was not significant in response times analyses ($F(1, 104) = 0.51$, $p = 0.475$, $\eta_p^2 < 0.01$), but was reliable in accuracy analyses ($F(1, 104) = 10.24$, $p = 0.002$, $\eta_p^2 = 0.09$), with higher accuracy for Facing ($M = 76.81\%$, SE = 1.22) compared to Non-Facing targets ($M = 74.69\%$, SE = 1.35).

### Experiment 2 (2500 ms) vs. Experiment 3 (1250 ms)

In an additional exploratory analysis, we directly compared the results of Experiments 2 and 3 in a 2 × 2 × 2 mixed ANOVA with a between-subjects factor of Presentation Time (2500 ms from Experiment 2; 1250 ms from Experiment 3), and within-subjects factors of Target Type (Facing or Non-Facing) and Set Size (4 or 8). First, this analysis revealed a main effect of Set Size ($F(1, 208) = 843.30$, $p < 0.001$, $\eta_p^2 = 0.80$), with faster response times for the smaller ($M = 1403.52$ ms, SE = 14.53) compared to the larger set size ($M = 1762.82$ ms, SE = 18.30). This main effect of Set Size was also qualified by a two-way interaction with Presentation Time ($F(1, 208) = 21.09$, $p < 0.001$, $\eta_p^2 = 0.09$), as the difference between set sizes was larger in Experiment 2 ($t(208) = 23.78$, $p < 0.001$, $d_z = 1.64$) compared to Experiment 3 ($t(208) = 17.29$, $p < 0.001$, $d_z = 1.19$). Most importantly, there was a reliable main effect of Target Type ($F(1, 208) = 11.07$, $p = 0.001$, $\eta_p^2 = 0.05$), with overall faster response times for Facing ($M = 1558.51$ ms, SE = 17.03) compared to Non-Facing targets ($M = 1607.83$ ms, SE = 17.00). This main effect was similar in the two experiments (no interaction between Presentation Time and Target Type, $F(1, 208) = 0.36$, $p = 0.548$, $\eta_p^2 < 0.01$), and in the two set sizes (no interaction between Presentation Time and Set Size, $F(1, 208) < 0.01$, $p = 0.977$, $\eta_p^2 < 0.01$). There were also no overall differences in inverse efficiency scores across the two experiments (no main effect of Presentation Time, $F(1, 208) = 1.03$, $p = 0.312$, $\eta_p^2 < 0.01$), and no three-way interaction between Presentation Time, Set Size, and Target Type ($F(1, 208) = 0.48$, $p = 0.491$, $\eta_p^2 < 0.01$).

### Effect of triad pose

In additional exploratory analyses, we examined any possible influence of target pose in a Target Type × Set Size × Target Pose repeated measures ANOVA, which revealed no interactions of the main effects of interest with Target Pose in Experiment 1 (all *F*s < 0.99, *p*s > 0.418) nor in Experiment 2 (all *F*s < 1.91, *p*s > 0.096) or Experiment 3 (all *F*s < 1.28, *p*s > 0.273).