

Social processing of dynamic naturalistic social interactions

Katie Daughters¹, Simona Skripkauskaite^{2,3}, Kami Koldewyn⁴

¹ Department of Psychology, University of Essex

² Department of Experimental Psychology, University of Oxford

³ Department of Psychiatry, University of Oxford

⁴ School of Human and Behavioural Sciences, Bangor University

Abstract

Research suggests that static depictions of social interactions preferentially capture our attention compared to non-interactions. Research also suggests that motion captures attention. To date, therefore, it is unknown whether *dynamic* social interactions preferentially capture attention relative to non-interactions, over and above motion cues. The present study captured 81 participants' eye-gaze and facial movements when viewing 4-second video clips of social-interactions compared to motion-matched non-interactions. We hypothesised that participants would spend more time looking at the two people in the videos relative to the background when viewing social interactions compared to non-interactions. Results confirmed our hypothesis and demonstrated that this effect was stronger for individuals with greater empathy and lower autistic traits. These results add to the growing body of research investigating the processing of social interactions in complex, naturalistic stimuli and demonstrate that social interactions do preferentially capture attention, even when motion cues are present.

Keywords: Social interaction, social attention, naturalistic, dyna

Introduction

Humans are an inherently social species. As such, researchers have sought to understand how we make sense of and engage with our social world. This has often meant focusing on one particular facet of social perception and conducting tightly controlled laboratory experiments using simplified stimuli. There is, however, a growing demand to move away from this approach to better understand how we process more complex and naturalistic social information, and in particular, whether this processing is specialised to or different for social interaction perception.

A large body of research has investigated visual attention as an important mechanism underlying social perception. It is known that social stimuli quickly and effectively capture attention relative to non-social stimuli (e.g., End & Gamer, 2017). More recently, studies have demonstrated a visual preference for 'interactive' or facing human dyads and triads compared to non-facing dyads and triads (Fratino et al., 2022; Papeo et al., 2019; Papeo & Abassi, 2019; Vestner et al., 2019, 2022). However, these studies typically use simple silhouettes and therefore do not encompass other important indicators to social interaction, such as eye-gaze or gestures.

A few studies have included more complex stimuli. For example, Villani et al. (2015) presented participants with paintings depicting two individuals either acting independently or interacting with one another. They found that participants spent more time looking at the bodies of individuals when they were interacting and more time looking at the face when acting independently. However, this study did not include a direct comparison of attention to social versus non-social information and thus it is not possible to ascertain the moderating impact of social interaction on this contrast; they also used paintings, not photographs depicting real life. A more recent study (Skripkauskaitė et al., 2022) expanded on this by investigating visual attention to interacting and non-interacting dyads in real life photographs. They found that participants spent more time looking at individuals compared to the background for interacting stimuli, and there was no difference for non-interacting stimuli, supporting the idea that social interactions, specifically, capture social attention.

These studies, however, have all utilised static images. Dynamic stimuli, on the other hand, not only enable the use of more ecologically valid stimuli but also incorporate additional indicators to interaction such as actions. Indeed, research demonstrates that movement also quickly and effectively captures attention (Atkinson et al., 2018). Introducing movement therefore provides an additional challenge as to whether social interactions capture attention differently or preferentially as previous studies would suggest. There is some preliminary evidence to support this hypothesis. For instance, observation of dynamic, compared to static, interactions give rise to different neural representation patterns (Landsiedel et al., 2022). Furthermore, Chevallier et al. (2015) found that children spent more time looking at faces compared to objects during dynamic videos that depicted multiple people and showed both social

interactions and non-interactions. This analysis, however, focused on faces exclusively, not whole individuals, and did not explicitly analyse differences between interactions and non-interactions. Thus, it remains unknown whether social interactions drive visual attention even in complex dynamic scenes where motion cues and other 'low level' visual features may compete for attention.

Visual attention is not the only mechanism through which we attempt to process and understand complex social information. As some of the previous studies suggest (Chevallier et al., 2015; Villani et al., 2015), the face is an important tool for navigating the social world. It can provide vital information, including affect, which is important when inferring the social context of an interaction. It is therefore important to correctly interpret facial cues and in particular other's emotions. Facial mimicry, the spontaneous copying of other's facial expressions, has been proposed to enhance emotion recognition (Hatfield et al., 1993; Stel et al., 2016) and empathy (Drimalla et al., 2019). It also facilitates social interactions through social learning (Kavanagh & Winkielman, 2016) and a desire to affiliate with others (Bourgeois & Hess, 2008). There is clear evidence of facial mimicry occurring when viewing another person's face in static (Dimberg et al., 2011) or dynamic (Daughters et al., 2021; Hess & Blairy, 2001) stimuli on a computer screen and when participating in live social interactions (Hess & Bourgeois, 2010; Stel et al., 2013). To the best of our knowledge, however, no study to date has examined facial movements in response to witnessing *others* in social interactions. The present study therefore aimed to fill this gap in the literature, to understand whether facial movements assist in the processing of complex social interactions.

Finally, it is important to consider individual differences that may bias one's visual attention and facial movements. Individuals who are higher in trait empathy have been found to be faster to look to human figures (Villani et al., 2015) and display greater facial movements (Dimberg et al., 2011) in response to static stimuli. In contrast, autistic individuals have been found to spend less time looking at human information compared to neurotypical individuals (e.g., Chevallier et al., 2015; Shi et al., 2015), while facial mimicry research has found both lower (Yoshimura et al., 2015) and similar (Van der Donck et al., 2021) facial movements in autistic compared to neurotypical individuals. Thus, although research suggests there are differences in both visual attention and facial movements for those with higher trait empathy and autistic traits, this is yet to be examined in naturalistic dynamic stimuli and has not been investigated with respect to social interaction perception.

The primary aim of the current study was to investigate visual attention and facial movements as two potential mechanisms underlying social perception of dynamic social scenes, and to assess to what extent the presence of a social interaction may moderate this effect. The secondary aim was to investigate the effect of individual differences on both visual attention to and facial movement in response to observed social interactions, contrasted with visually similar non-interactions. This was examined across four hypotheses: 1) Participants will spend more

time looking at social (versus non-social) information and this will be moderated by the presence of social interactions, such that participants will spend more time looking at social (versus non-social) information in social interactions compared to non-interactions; 2) Participants will display more facial movement (specifically, frowning) in response to social interactions, compared to non-interactions; 3) Participants higher in trait empathy will display greater visual attention and facial movements in response to social interactions (compared to non-interactions); 4) Participants lower in autistic traits will display greater visual attention and facial movements in response to social interactions (compared to non-interactions).

Method

Participants

An a priori power calculation determined that 70 participants would be sufficient to attain 80% power with an $\alpha = 0.05$ for a three-way interaction, however, due to anticipated data loss, the study aimed to recruit 85 participants. Eighty-one participants (19-84, $M = 30.78$, $SD = 13.85$; 45 females) took part in the in-person study at ESSEX Labs. Participants were recruited via an online booking system and received financial compensation for taking part. The data were pre-processed and analysed in accordance with pre-registration (https://aspredicted.org/blind.php?x=VG_VY4C). Data from seven participants were removed due to true outliers (e.g., technical issues); none of the participants were excluded due to missing eye-tracking data or missing facial analysis data; 15 participants had missing questionnaire data and were therefore not included in the individual differences analyses. The study was approved by the University of Essex Department of Psychology Ethics Committee (ETH2122-1116).

Materials

Social Perception Task

Participants were asked to watch 60 video clips depicting various everyday scenarios. Each 4-second video clip depicted a scenario with two individuals either engaging in a social interaction or acting independently (non-interaction). In total, there were 30 interaction and 30 non-interaction scenarios where each interaction scene was matched as closely as possible to a non-interaction scene, such that they both contained the same people and objects, and where people performed similar actions and (where possible) maintained the same physical distance between the two individuals (e.g., Figure 1). Matched interaction and non-interaction scenes were filmed from the same camera angle in the same setting and did not differ in overall motion energy (see Landsiedel et al., 2022 for analysis). Interaction videos were previously subjectively rated as more social, more positive, and more visually interesting than non-interaction videos (Landsiedel et al., 2022). Scenarios were acted out by four different actor pairs (two female/female

pairs, one male/male pair and one female/male pair) captured in eight different locations (e.g., an office, lobby waiting area etc.).

The experimental task was programmed using iMotions (www.imotions.com), a biometric research platform that can be used to synchronize multiple psychophysical measures. This enabled automated facial coding, eye tracking, and stimulus presentation to be precisely coordinated. Eye-tracking data was recorded using a portable Tobii x2-120 compact eye-tracker sampling at 120Hz with a screen resolution of 1920x1080. An I-VT fixation filter was applied, and data were sampled from both eyes to produce information on eye position and latency. Participants' facial movements were recorded via a Logitech HD webcam and post-processed using the AFFDEX algorithm for automatic facial coding developed by Affectiva Inc. (El Kaliouby & Robinson, 2005; McDuff et al., 2010).

Questionnaires

Autistic Traits. The Autism Quotient Short version (AQ-S; Kloosterman et al., 2011) is an adapted 28-item version of the original 50-item Autism Quotient (Baron-Cohen et al., 2001). Items relate to five subscales: social skills, mind reading/communication, restricted and repetitive behaviour, imagination and attention to detail. Participants were asked to rate to what extent they agree/disagree with each item (1 = "Definitely Agree"; 4 = "Definitely Disagree"). The measure includes reverse item scoring such that higher scores represent more autistic traits. A total score and mean scores for each subscale were calculated. The total score achieved good internal reliability (Cronbach's $\alpha = 0.78$).

Trait Empathy. The Interpersonal Reactivity Index (IRI; Davis, 1983) is an established questionnaire, with items pertaining to four subscales: empathic concern, fantasy, personal distress and perspective taking. There are 28 items in total, seven for each subscale (including reverse coded items). For each item, participants are asked to indicate on a 5-point scale to what extent the statement can be applied to them (1 = "does not describe me very well"; 5 = "describes me very well"). A total score and mean scores for each subscale were calculated; higher scores indicated greater empathy. The total score achieved good internal reliability (Cronbach's $\alpha = 0.86$).

Procedure

Small groups (between 5-10) of participants took part in the study in a large testing room, with individual participants spaced around the room in private computer cubicles. The blinds were drawn and the same lighting was used for each testing session. Before data collection, participants read the study information sheet and provided written informed consent. Participants were then seated 60-65cm from the screen and completed a 9-point calibration. If the calibration quality was poor, the process was repeated. No participant was asked to complete the calibration more than three times. They then completed both questionnaires via the online survey software, Qualtrics, before starting

the social perception task. Participants were instructed to watch the videos in a free viewing task. Each of the 60 trials consisted of a grey background with a black fixation cross presented in the middle of the screen for 10 seconds¹ before the 4-second video. The videos began automatically or were manually triggered earlier by the participants, who were instructed to focus on the fixation cross before beginning each video. The videos were presented in random order and occupied the full screen. Participants' eye and facial movements were recorded throughout the task. At the end of the study, participants were fully debriefed before leaving.

Data Analysis

All statistical analyses were conducted in R (version 1.4.1717) using the lme4 package (Bates et al., 2015). Satterthwaite's approximate method was used for significance testing (lmerTest package; Kuznetsova et al., 2017). The data and analysis code are available on OSF (https://osf.io/dqjpr/?view_only=01bdb55be14148e4ae4121b5f64489fa).

Eye tracking

Dynamic Areas of Interest (AOIs) were drawn around the two individuals in each video (see Figure 1) using iMotions. These AOIs were adjusted on a frame-by-frame basis to ensure that all the human figure was captured in each frame. A static AOI was also created for each video capturing the whole scene. Dwell times (time the participants were fixated in an AOI while the stimulus was on the screen) was then exported for all AOIs. Thus, by subtracting the dwell time for both individuals from the scene AOI we could create a measure of time spent looking at the background or non-social AOI in each video.

Linear mixed-effect (multilevel) modelling was carried out to assess whether participants spent more time looking at social AOIs compared to non-social AOIs, and whether looking time differences were moderated by the presence of a social interaction. Thus, modelling focused on a 2(Stimulus: interacting=1 vs. non-interacting=0) x 2(AOI: social=1 vs. non-social=0) design. Working from the null/unconditional model, then adding random then fixed effects, models assessed the impact of individual differences in empathy and autistic traits. Participants were modelled as a random intercept effect; stimulus, AOI, IRI scores and AQ scores as fixed predictors.

Facial movements

Participants' facial movements were scored in terms of 20 'channels' (Ekman & Friesen, 1975) based on the Facial Action Coding System (FACS) action units in AFFDEX. Scores ranged from 0 to 100 representing the probability that that channel was being expressed (0 - channel not expressed, 100 - channel expressed). Because AFFDEX offers a large number of output metrics, to avoid type 1 errors due to multiple tests, careful consideration is required to identify relevant metrics in advance. In line with our pre-registration, metrics for the

¹ This was due to an iMotions programming limitation of a minimum presentation time of 10 seconds.

‘smile’ and ‘brow furrow’ AFFDEX channels for each video were exported for analysis. However, the smile channel did not produce enough data within reliable levels of activation (less than 1% after thresholding) to warrant further investigation. Therefore, no analysis will be reported for smiling. There was more variation in activation across the frown channel (5% after thresholding), however, this was heavily skewed such that most participants displayed very little frowning across the study. Therefore, in line with iMotions recommendations and to avoid significant skewness in the data, a threshold of 25 was applied to the data (Daughters et al., 2021), such that any datapoints above 25 were deemed to represent that channel being expressed; anything below 25 and the channel was deemed not to be expressed.

Logistic linear mixed-effect (multilevel) modelling was carried out to assess whether participants displayed more frowning during social interactions compared to non-interactions. Due to the dynamic nature of the stimuli, and that social interaction typically takes place in the middle of the 4-second videos, time was also added as a predictive factor into the analysis. Thus, these models focused on 2(Stimulus: interacting=1 vs. non-interacting=0) x 4(Time:0-3 second) design. Models also assessed the impact of individual differences in empathy and autistic traits. Participants and video identifiers were modelled as crossed random effects (because we anticipated that there would be some random variation in valence/affective tone across videos which might impact frowning); stimulus, AOI, IRI scores and AQ scores as fixed predictors. IRI and AQ scores were centred and scaled, however, all models including these factors failed to converge and will therefore not be reported.

Results

Eye Tracking

Both stimulus and AOI significantly predicted dwell time (see Table 1 for statistics), however, there was also a significant interaction in the anticipated direction. Participants spent more time looking at social AOIs (compared to non-social AOIs) when viewing social interactions (compared to non-interactions); see Figure 2.

The model achieved an overall better fit, however, when individual differences in empathy and autistic traits were added in. Again, AOI significantly predicted participants dwell time, but there were also significant interactions between AOI and both IRI and AQ scores (see Table 2 for statistics). Participants with greater trait empathy spent longer looking at social AOIs (see Figure 3) and participants with more autistic traits spent longer looking at non-social AOIs. Indeed, participants with particularly high autistic traits in our sample crossed over to spend more time looking at non-social AOIs than social (see Figure 4). These trait related effects, however, were not further moderated by the stimulus type.

Facial Movements: Frowning

Both stimulus and time bin significantly predicted the presence of frowning (see Table 3 for statistics). There was also a significant interaction. In contrast to our hypothesis, participants displayed more frowning during non-interactions compared to interactions, an effect that was very slightly more pronounced during the first three seconds of the videos (see Figure 5). However, this effect was small and all the fixed effects in the model explained only 0.2% of frowning variance.

Discussion

The present study investigated two potential mechanisms that might facilitate social interaction perception: visual attention and facial movements. Crucially, the study sought to extend existing research by using *dynamic naturalistic* stimuli, thereby examining whether visual attention, in particular, is still captured preferentially by social interactions even when motion cues – which also drive visual attention – are present. We also sought to explore the effect of social interactions on facial movements when witnessing “everyday” dyadic interactions. The findings support our first hypothesis; participants spent longer looking at the two individuals in the video relative to the background when they were interacting compared to not interacting. In addition, and partially supporting hypotheses three and four, there were also individual differences in the extent to which social information captured attention. Higher scores in trait empathy were associated with increased social bias (i.e., more time spent longer looking at the people in the video and less time looking at the background). In contrast, greater levels of autistic traits were associated with reduced social bias. Interestingly, there was a cut-off point on the AQ-S where individuals’ preference for looking swapped from people to the background (a score of approximately 76-77 in our sample). However, contrary to our expectations, there was no significant interaction between either empathy or autistic traits and stimulus type, suggesting that these differences in social bias were not moderated by the presence of social interactions.

Our findings support and extend the small body of research investigating the visual processing of social interactions (e.g., Chevallier et al., 2015; Fratino et al., 2022; Papeo & Abassi, 2019; Skripkauskaite et al., 2022; Vestner et al., 2019). Specifically, we replicate an earlier finding that social interactions capture visual attention to a greater extent than non-interactions in naturalistic stimuli (Skripkauskaite et al., 2022), but extend this finding by demonstrating this in *dynamic videos* that may provide both additional cues to interaction (e.g., coordinated or reciprocal actions) but also greater distraction (due to ‘low level’ motion cues capturing attention) in the non-interaction condition. We also replicate and extend findings from the only study to date to assess visual attention to social (faces) versus non-social (objects) AOIs in naturalistic dynamic stimuli (Chevallier et al., 2015). In their study, participants demonstrated greater looking to faces versus objects, and thus our findings 1) confirm that whole bodies also capture visual attention to a greater extent than background objects in naturalistic dynamic stimuli, and 2) extend this by demonstrating this effect to be moderated by the

presence of a social interaction. We note that Chevallier et al. (2015) did include interaction versus non-interaction conditions in their task. However, as this factor was not included in their analysis, it is not possible to draw any conclusions regarding social interaction perception from their study.

Why might individuals spend more time looking at two individuals when they are interacting as opposed to when they are engaged in pursuing separate goals, particularly when scenes contain the same number of people? On the surface, scenes where agents pursue separate goals might require *more* attention and processing to parse than scenes where individuals engage in joint actions. However, as discussed in prior work (e.g., Skripkauskaitė et al., 2022) social interactions constitute more than the sum of their parts; for example, they provide information on the social context and the relationship between two individuals. In addition, and as a potential explanation for why we observed greater visual attention to interacting people in *dynamic stimuli*, as the interaction unfolds over time individuals may spend more time looking back and forth between the interactants (compared to non-interacting individuals). Future research may wish to investigate this hypothesis using scan path analysis. Lastly, people find interactive scenes – even when they are low-emotive everyday scenes – to be more interesting and more positive than non-interactive scenes, even when these scenes are visually similar, matched for motion energy, and include the same people performing similar actions. In our view, this is a feature of social interactions rather than a design flaw. Social interactions may draw visual attention *because* humans are inherently social and pursue understanding interactions and relationships even when doing so serves no obvious purpose for the observer.

Our findings are also broadly in line with previous research investigating individual differences in visual attention. While Villani et al. (2015) found that individuals with greater empathic concern (a subscale of the IRI) were faster (fewer fixations prior) to look to the face, they found no effect of any IRI subscale on looking time. In contrast, the present study found that individuals higher in trait empathy (represented by the whole IRI scale) spent longer looking at people. The findings across studies are broadly compatible - individuals with higher empathy show a stronger attentional bias towards social information – and differences between studies are potentially explained by methodological differences. Villani et al. presented static images of artwork for 15 seconds, while the current study presented naturalistic dynamic videos for 4 seconds. First, time-to-first-fixation or similar metrics are not relevant in stimuli that unfold over time and thus were not calculated in the current study. Secondly, the longer presentation time used by Villani and colleagues may have resulted in general processing of the image towards the end, potentially masking possible biases in looking time that might have been present during early processing. Chevallier et al. (2015)'s primary interest was to examine differences in visual processing of social information between autistic and non-autistic individuals. Indeed, and in line with our findings, they found that autistic individuals

spent less time looking at social AOIs and more time looking at object AOIs. Thus, our results confirm that individuals with higher levels of autistic traits also display differences in social information processing in naturalistic dynamic stimuli. We note, however, that our individual difference findings were not moderated by the presence of social interactions. This may reflect a genuine finding, suggesting that individual differences could explain a significant portion of variance in our main finding. However, we suggest this is unlikely due to the within subject design. Our study was powered for a three way rather than a 4-way interaction, and so it is possible that we may not have had enough power to detect higher order effects in our more complex model incorporating individual differences. Future work, perhaps manipulating stimulus type to include stimuli that induce empathic responses, could delve more deeply into individual differences that might specifically influence attention specifically to social interactions.

The available data for facial movements did not support our second hypothesis: participants displayed more frowning in response to non-interactions compared to interactions, albeit this effect was very small. It may be that rather than participants utilising their own facial movements to make sense of the social interactions (Bourgeois & Hess, 2008; Drimalla et al., 2019; Kavanagh & Winkielman, 2016), some participants were employing greater facial movements, in particular frowning, during non-interactions in an effort to make sense of the information (i.e., they were searching for social interactions). We acknowledge, however, that our results are based on a low (25% probability) threshold and very small effect sizes and thus caution is advised when interpreting the results. Unfortunately, models looking at facial movements incorporating trait empathy and autistic traits failed to converge and thus no conclusions can be drawn about the possible impact of individual differences on facial activity in response to witnessing social interactions. We believe the strongest argument for why we failed to observe reasonable levels of facial movements in the current study is due to the webcam-based iMotions algorithm, which may not be sensitive enough to detect the subtle facial movements elicited by the stimuli. Indeed, the authors have encountered this issue in a previous study in which participants were ultimately deliberately instructed to mimic the study stimuli in order to achieve reliable measures of facial activity (Daughters et al., 2021). Future research could investigate the impact of social interactions on facial movements using more sensitive measurement techniques such as facial electromyography and/or include social scenes with more emotive content in the stimulus set.

Although failing to acquire adequate facial data is a significant limitation of the current study, we highlight several strengths as well. First, the study recruited a good sample size ($N = 85$, eye tracking observations = 7018) and utilised mixed-effect modelling to control for random variation across participants. Second, the study used naturalistic dynamic videos and, as such, provides novel insight into the processing of social interactions in more ecologically valid stimuli. Third, the stimulus set, although naturalistic, is tightly controlled such that

important factors such as proximity between agents, props, actions, and level of motion is equivalent across specific interaction and non-interaction contexts. Fourth, although the mean age (30.78) of the current study is close to the typical student populations utilised in many psychological experiments, our age range (19-84) incorporated a much wider proportion of the population and achieved a roughly 50/50 sex split, thereby increasing the generalisability and inclusivity of the study.

In conclusion, the present study investigated two potential mechanisms through which individuals may process social interactions. Importantly, the current study used naturalistic dynamic stimuli and therefore provides novel insights into social interaction processing. The study found preferential processing of people in social interactions versus non-interactions, measured via longer looking times. There were also individual differences in this processing, such that individuals higher in trait empathy spent longer looking at people relative to the background, while individuals higher in autistic traits spent longer looking at the background. Future studies may wish to explore whether these longer looking times are driven by continual looking back and forth between interactants as the social interaction unfolds over time.

Declaration of Conflicting Interests

The authors declare that there is no conflict of interest.

Funding

This work has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC-2016 -STG-716974: Becoming Social).

References

- Atkinson, M. A., Simpson, A. A., & Cole, G. G. (2018). Visual attention and action: How cueing, direct mapping, and social interactions drive orienting. *Psychonomic Bulletin & Review*, 25, 1585–1605.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, 77(3), 343–352.
- Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, 8(5), 620–628.
- Daughters, K., Manstead, A. S. R., & Van Der Schalk, J. (2021). Oxytocin and emotion recognition: Investigating the possible roles of

- instructed mimicry and eye gaze. *Current Research in Ecological and Social Psychology*, 2, 100019.
- Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113.
- Dimberg, U., Andréasson, P., & Thunberg, M. (2011). Emotional empathy and facial reactions to facial expressions. *Journal of Psychophysiology*, 25(1), 26–31.
- Drimalla, H., Landwehr, N., Hess, U., & Dziobek, I. (2019). From face to face: the contribution of facial mimicry to cognitive and emotional empathy. *Cognition and Emotion*, 33(8), 1672–1686.
- Ekman, P., & Friesen, W. V. (1975). Measuring facial movement. In *Pictures of facial affect* (Vol. 1, Issue 1). Consulting Psychologists Press. <https://doi.org/10.1007/BF01115465>
- El Kaliouby, R., & Robinson, P. (2005). Real-time inference of complex mental states from facial expressions and head gestures. In *Real-time vision for human-computer interaction* (pp. 181–200). Springer.
- End, A., & Gamer, M. (2017). Preferential processing of social features and their interplay with physical saliency in complex naturalistic scenes. *Frontiers in Psychology*, 8, 418.
- Fratino, V., Colombatto, C., & Ristic, J. (2022). From dyads to triads: Perceptual unity of social groups. *Journal of Vision*, 22(14), 4227.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, 2(3), 96–100.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, 40(2), 129–141. [https://doi.org/10.1016/S0167-8760\(00\)00161-6](https://doi.org/10.1016/S0167-8760(00)00161-6)
- Hess, U., & Bourgeois, P. (2010). You smile–I smile: Emotion expression in social interaction. *Biological Psychology*, 84(3), 514–520.
- Kavanagh, L. C., & Winkielman, P. (2016). The functionality of spontaneous mimicry and its influences on affiliation: An implicit socialization account. *Frontiers in Psychology*, 7, 458.
- Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. A. (2011). Evaluation of the factor structure of the Autism-Spectrum Quotient. *Personality and Individual Differences*, 50(2), 310–314.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Landsiedel, J., Daughters, K., Downing, P. E., & Koldewyn, K. (2022). The role of motion in the neural representation of social interactions in the posterior temporal cortex. *NeuroImage*, 262, 119533.
- McDuff, D., El Kaliouby, R., Kassam, K., & Picard, R. (2010). Affect valence inference from facial action unit spectrograms. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, 17–24.
- Papeo, L., & Abassi, E. (2019). Seeing social events: The visual specialization for dyadic human–human interactions. *Journal of*

Experimental Psychology: Human Perception and Performance, 45(7), 877-888.

- Papeo, L., Goupil, N., & Soto-Faraco, S. (2019). Visual search for people among people. *Psychological Science*, 30(10), 1483-1496.
- Shi, L., Zhou, Y., Ou, J., Gong, J., Wang, S., Cui, X., Lyu, H., Zhao, J., & Luo, X. (2015). Different visual preference patterns in response to simple and complex dynamic social stimuli in preschool-aged children with autism spectrum disorders. *PloS One*, 10(3), e0122280.
- Skripkauskaitė, S., Mihai, I., & Koldewyn, K. (2022). Attentional bias towards social interactions during viewing of naturalistic scenes. *Quarterly Journal of Experimental Psychology*, 17470218221140880.
- Stel, M., Hess, U., & Fischer, A. (2016). The role of mimicry in understanding the emotions of others. In U. Hess & A. Fischer (Eds.), *Emotional Mimicry in Social Context* (pp. 27-43). University Press.
- Stel, M., van den Bos, K., Sim, S., & Rispens, S. (2013). Mimicry and just world beliefs: Mimicking makes men view the world as more personally just. *British Journal of Social Psychology*, 52(3), 397-411.
- Van der Donck, S., Vettori, S., Dzhelyova, M., Mahdi, S. S., Claes, P., Steyaert, J., & Boets, B. (2021). Investigating automatic emotion processing in boys with autism via eye tracking and facial mimicry recordings. *Autism Research*, 14(7), 1404-1420.
- Vestner, T., Over, H., Gray, K. L. H., & Cook, R. (2022). Objects that direct visuospatial attention produce the search advantage for facing dyads. *Journal of Experimental Psychology: General*, 151(1), 161.
- 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.
- Villani, D., Morganti, F., Cipresso, P., Ruggi, S., Riva, G., & Gilli, G. (2015). Visual exploration patterns of human figures in action: an eye tracker study with art paintings. *Frontiers in Psychology*, 6, 1636.
- Yoshimura, S., Sato, W., Uono, S., & Toichi, M. (2015). Impaired overt facial mimicry in response to dynamic facial expressions in high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 45, 1318-1328.