Check for updates

# Gaze behavior in response to affect during natural social interactions

Antonia Vehlen[1], Artem V. Belopolsky[2] and Gregor Domes[1,3]*

[1]Department of Biological and Clinical Psychology, University of Trier, Trier, Germany, [2]Department of Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [3]Institute for Cognitive and Affective Neuroscience, University of Trier, Trier, Germany

Attention in social interactions is directed by social cues such as the face or eye region of an interaction partner. Several factors that influence these attentional biases have been identified in the past. However, most findings are based on paradigms with static stimuli and no interaction potential. Therefore, the current study investigated the influence of one of these factors, namely facial affect in natural social interactions using an evaluated eye-tracking setup. In a sample of 35 female participants, we examined how individuals' gaze behavior responds to changes in the facial affect of an interaction partner trained in affect modulation. Our goal was to analyze the effects on attention to facial features and to investigate their temporal dynamics in a natural social interaction. The study results, obtained from both aggregated and dynamic analyses, indicate that facial affect has only subtle influences on gaze behavior during social interactions. In a sample with high measurement precision, these findings highlight the difficulties of capturing the subtleties of social attention in more naturalistic settings. The methodology used in this study serves as a foundation for future research on social attention differences in more ecologically valid scenarios.

KEYWORDS

gaze behavior, social interaction, facial affect, dynamic analysis, eye-tracking

## 1 Introduction

Interactions are a major part of human social life. Success relies on allocating attention to key social cues (i.e., diagnostic cues). In social interactions cues become diagnostic, which allow conclusions to be drawn about the identity of the interaction partner, his or her feelings, the direction of attention, and the content of what is being said. The face, and in particular the eye region of the interaction partner, are a rich source of social information and therefore attract the most attention (Birmingham et al., 2009; Devue et al., 2012; Yarbus, 1967). Eye movements are closely tied to the cognitive and emotional processes underlying social interactions, as proposed by the *eye-mind hypothesis*, which suggests a direct link between where we look and what we process cognitively (Just and Carpenter, 1984). Retrieving social information is facilitated when the partner responds with direct gaze (Senju and Johnson, 2009). Apart from information retrieval, eye contact has the function of conveying information to the interaction partner, e.g., about one's own current mental state or intended further actions (*dual function of gaze*; Argyle and Cook, 1976; Risko et al., 2016). Despite a general preference for faces and the eye region, studies reveal substantial variance in social attention.

The recording of gaze behavior in standardized setups was used as an indicator of attention allocation (Wright and Ward, 2008), and systematic experimental manipulations helped identify factors contributing to the variance in attentional preferences during face

viewing (for review see: Dalmaso et al., 2020; Hadders-Algra, 2022). On the observer side, factors such as age, cultural background, and potential psychopathologies have been shown to influence these processes. While attention to faces in general seems to decrease with age (De Lillo et al., 2021), individuals from Western cultures have been found to fixate more on the mouth region compared to individuals from Asian cultures (Senju et al., 2013). Furthermore, altered processing of the eye region has been linked to several pathologies such as schizophrenia (Loughland et al., 2002), autism spectrum disorder (Itier and Batty, 2009; Setien-Ramos et al., 2022), and social anxiety (Chen et al., 2020). It is believed, that altered attentional processing contribute to impaired social functioning in these disorders.

In addition to observer characteristics, variance in attentional preferences can also be explained by factors inherent to the stimulus (e.g., interaction partner). One factor that has been studied continuously is the affective state of the stimulus. Affective states, such as those mediated by emotional expressions, have been linked to attentional preferences, as, for example, the mouth region attracts more attention when positive affect, such as a happy stimulus, is viewed (Beaudry et al., 2014; Calvo et al., 2018; Green et al., 2003). Findings are less clear for negative stimuli and can vary depending on factors such as affect type and study design. Some studies have reported an increase in attention to the eye region for sad and angry stimuli (Beaudry et al., 2014; Calvo et al., 2018; Eisenbarth and Alpers, 2011), while others have observed a decrease in attention for fearful and angry stimuli (Hunnius et al., 2011). These heterogenous findings indicate the complexity and variability of the relationship between affect and gaze behavior.

Although these studies have provided valuable insights into information retrieval through gaze behavior, the standardization used limited the setups to computer screens and thus never captured the signaling function of gaze. The impact of this simplification became clear through a series of studies applying a *waiting room paradigm* (Grossman et al., 2019; Horn et al., 2022; Kulke et al., 2023; Laidlaw et al., 2011; Rösler et al., 2021). In these studies, it was demonstrated that attention to faces varied significantly depending on the social context (e.g. screen vs. live). With advancing technology, the field has developed more naturalistic setups that capture the dynamics of gaze behavior, allowing for greater complexity in the study of social attention (Hessels et al., 2017; Tönsing et al., 2022). However, such naturalistic studies are very challenging to conduct, as they require extensive knowledge of the factors that influence data quality and the possibilities of standardization and validation (Nebe et al., 2023; Vehlen et al., 2021, 2022). In addition to the technical challenges, data collection in these naturalistic studies is often so complex that it requires extensive training for the individuals involved. In order to ensure that observed behavior in social interactions is not influenced by personal factors, experimenters and research assistants must learn to modulate their behavior as consistently and convincingly as possible. Despite these challenges, recent reviews emphasize the importance of integrating the historical best practices of eye-tracking with these modern methodologies, advancing the understanding of gaze functions in complex and interactive environments (Carter and Luke, 2020; Hessels, 2020).

The present study applied such a setup to investigate gaze behavior in natural social interactions, as evaluated by Vehlen et al. (2021). The setup was used to observe gaze dynamics in response to facial affect modulations by an interaction partner. More specifically, participants interacted with trained research assistants (RAs) who were instructed to show brief sequences of positive and negative facial affect at defined time points during a structured conversation, while participants' gaze behavior was recorded using remote eye-tracking. Based on the above literature of standardized setups, we hypothesize that there will be significant differences in the distribution of attention to the eye and mouth regions depending on the observed facial affect. Specifically, we expect that positive facial expressions will result in greater attention to the mouth region, negative expressions will lead to greater attention to the eye region. Furthermore, we hypothesize that these attentional patterns will exhibit time-dependent variability, reflecting the dynamics of natural social interactions. Specifically, we predict that differences in attention distribution will be most pronounced during the affect modulation period, rather than before or after. To capture these nuances, we have employed analytical methods that allow us to observe and quantify these dynamic effects over time. The longer-term goal is to contribute to our understanding of successful interactions, with potential applications in fields such as psychotherapy or robotics (Hessels, 2020).

## 2 Methods

### 2.1 Participants

Fifty-one female participants took part in the study. They were recruited via internal communications of the University of Trier and via flyers distributed on campus. Exclusion criteria consisted of visual impairments exceeding $\pm$ 1.5 diopters, any type of prescriptive visual aid other than soft contact lenses, acute eye inflammation and prior eye surgery. In the non-clinical sample psychotherapeutic treatment in the last 2 years, neurological diseases, and residual symptoms of a brain injury or the intake of psychotropic drugs also led to exclusion of participants. The sample was limited to heterosexual female participants to reduce attraction effects in interactions with female RAs. Six participants were excluded from analysis because of technical problems with the eye-tracker during testing sessions ($n = 6$). Four additional participants were excluded because the facial affect modulation was not performed, as reported by the research assistant who did not receive signals through the headphones ($n = 4$). Three participants were excluded because they commented on the affect modulation performed by the RAs ($n = 3$). During the session, other participants retrospectively reported watery eyes ($n = 1$), visual impairment above the predefined threshold ($n = 1$) and astigmatism ($n = 1$), which also led to exclusion. The final sample consisted of 35 women with a mean age of 22.3 $\pm$ 3.1 years (range: 18–29 years).

The study was conducted in accordance to the Declaration of Helsinki and was approved by the ethics committee of the University of Trier. All participants were informed about the study, gave written informed consent, and received compensation of €10.

The pictures on which a female person is depicted are used with the permission of this person, who has given her informed consent for publication.

## 2.2. Setup, calibration and validation

Gaze data was recorded at a sampling frequency of 120 Hz with a Tobii X3-120 remote infrared eye-tracker and the Tobii Studio Pro software (version 3.4.5). The eye-tracker was placed on a square table using a fixed monopod on which it was mounted at a 28° angle. The scene was recorded with a webcam (Logitech c920; 30 fps, 1,920 × 1,080 px resolution) attached above the heads of the participants. The operating distance (distance between the participant's eyes and the eye-tracker) was set at 65 cm, while the viewing distance (distance between the eyes of the interaction partners) was constant at 131 cm (detailed description and an illustration of the setup with the exact distances can be found here: Vehlen et al., 2021). The nine-point calibration sheet was attached to a partition wall and covered an area of 40 × 40 cm (21.6 × 21.6° visual angle) with a inter-point distance of 20 cm (∼8.7° visual angle). The sheet was placed approximately at the height at which the face of the RA was expected to be. The calibration was followed by three validation sequences over the course of the entire testing session. The first validation sequence immediately followed the calibration by repeating five of the calibration points (except for the points in the four corners) in random order (*wall validation*). Another validation sequence took place just before the interaction by instructing the participant to fixate facial features of the interaction partner (left eye, right eye, nose, and mouth; *pre-validation*). Since the facial features are moving target points, a frame-wise definition of the corresponding coordinates had to be performed. This was possible with the facial recognition software OpenFace (Amos et al., 2016), from whose facial landmarks the target points were calculated. The interaction was followed by another validation sequence on the face of the interaction partner (*post-validation*). Data quality was also estimated during the interaction and correlated with the movement of the interaction partner. These results can be considered in more detail in the evaluation study of the setup (Vehlen et al., 2021) and are not included in this paper.

## 2.3. Face-to-face interaction

For the face-to-face interaction, two RAs were trained to modulate their facial affect in response to instructions transmitted through in-ear headphones. They were taught to modulate their facial expressions to express negative, neutral, and positive affect. During the actual interactions the background (gray curtains) and the clothing of the RAs (dark without prints) were kept constant. The interaction sequences were structured by 12 questions from a published paradigm (Aron et al., 1997). The questions were alternately read aloud by the RA and the participant. Both the RA and the participant answered each question, resulting in two responses for each question. The questions were designed to facilitate interpersonal connection and encompassed a diverse

array of topics. Examples included selecting a dinner guest from any individual in the world, discussing aspirations for fame, and reflecting on personal habits and life narratives. The full list of questions can be found in the Supplementary material. To ensure consistency and promote a positive interaction atmosphere, the RAs prepared their responses in advance. Prior to providing their answers, both the RA and the participant would press a movable buzzer on the table. When the participant pressed the buzzer and answered the question, the buzzer triggered a Raspberry Pi running an internal script that sent the instructions (e.g., "Look angry") to the RAs' in-ear headphones (for details see Figure 1A). The whole interaction lasted on average 17.31 ± 3.36 min. The RAs showed each affect types (negative, neutral, and positive) four times during the interaction while participants were responding to the questions. The order of affect types was randomized.

## 2.4. Procedure

After an online screening that assessed the exclusion criteria, participants were invited to the laboratory. Each testing session began with the calibration of the RA, who was presented to the participants as an unknown interaction partner. However, in this study, only the participants' gaze data were analyzed. Subsequently, the participant was calibrated. Each calibration was followed by the first validation sequence. When the calibration wall was replaced with the RA's chair, both interaction partners, now seated across from each other, performed another validation sequence on the face of the interaction partner. The social interaction then took place, followed by another round of face validation. Last, participants completed questionnaires that captured constructs such as social anxiety, gaze anxiety and autistic traits. Summary statistics of these measures are provided in Supplementary Table S1.

## 2.5. Data analysis

### 2.5.1. Preprocessing

All calculations of gaze data were based on averaged binocular data and the total dwell times were chosen as gaze measure. In total dwell times all gaze behavior (fixations, saccades, etc.) directed at a given region within a given time period are analyzed.

### 2.5.2. Robustness, precision, accuracy

Robustness or trackability is a measure of the amount of data lost due to failed pupil detection, blinks of the participant or looks outside the tracking area. Lower values correspond to higher data loss. The accuracy can be understood as a measure of the validity of the gaze data and is high if the gaze coordinates recorded by the eye-tracker match the target point of the observer. Thus, low values correspond to high accuracy. Precision reflects the reliability of the gaze data and is high when continuous gaze at a given point in space results in gaze coordinates with a low dispersion. Again, low values correspond to high precision. More detailed information on the calculation of these three measures can be found here (Vehlen et al., 2021).
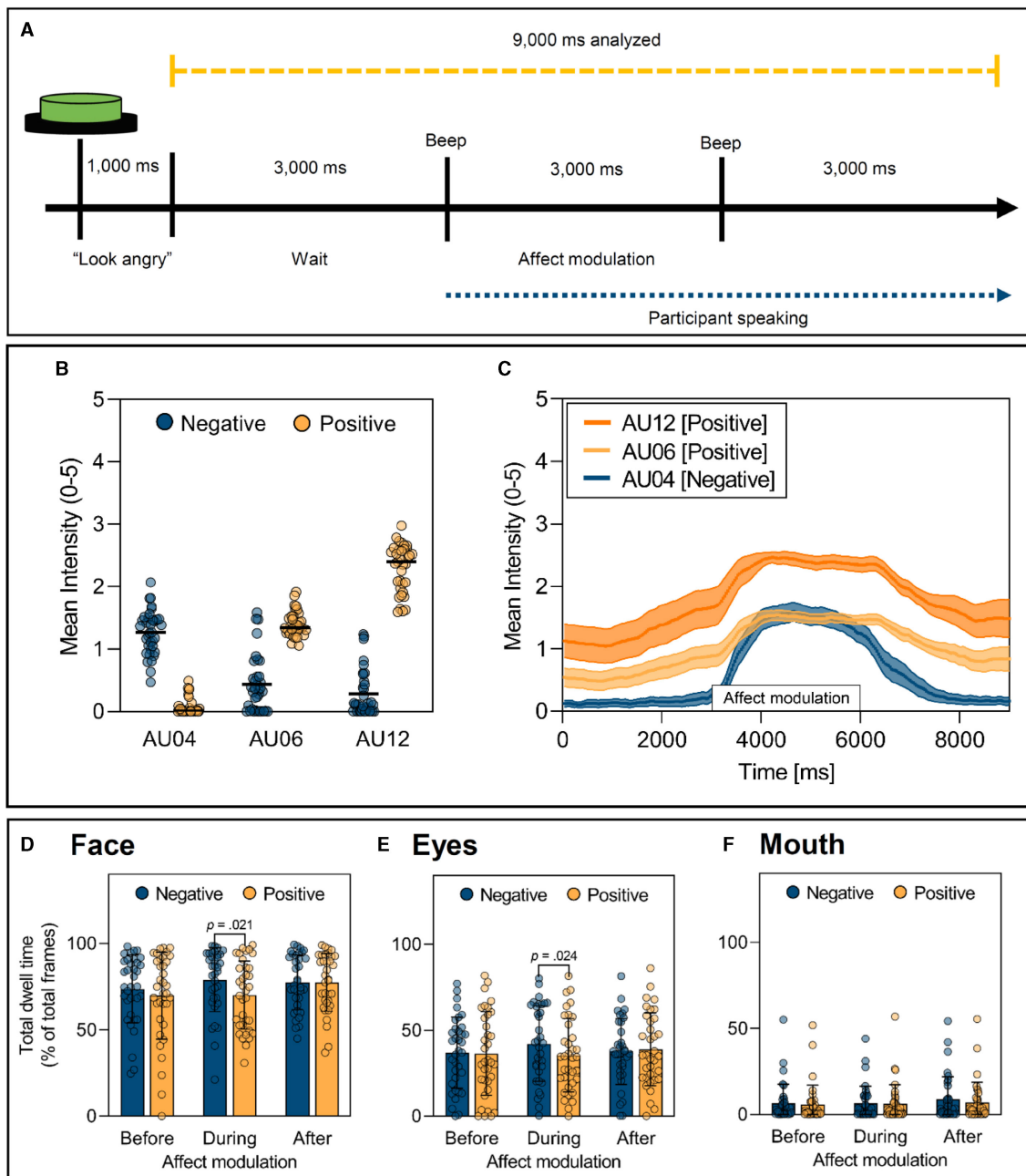
FIGURE 1
Procedure and effect of facial affect modulation within natural social interactions. (A) Timing of affect modulation. Instructions (e.g., "Look angry") are transmitted through in-ear headphones worn by the Research Assistant (RA). The blue line indicates the time during which the participant speaks. The yellow line illustrates the data analyzed in the study. (B, C) Activation of Action Units (AUs) during affect modulation. (B) Mean intensity of AU activation during the three seconds of positive and negative affect averaged over all trials. Black lines represent the mean intensities over all participants. (C) Mean intensity of AU04, AU06 and AU12 activation over time (9 seconds) averaged over all trials and participants. Solid lines represent mean values with corresponding 95%-confidence intervals visible by the shaded areas. (D–F) Gaze on facial Areas Of Interest (AOIs) as a function of affect type and time. (D) Summed total dwell times on all facial AOIs. (E) Total dwell times on the eyes. (F) Total dwell times on the mouth. Before = 0–3,000 ms; During = 3,000–6,000 ms; After = 6,000–9,000 ms. Bars represent $M \pm SD$.

## 2.5.3. Validation of facial affect modulation

To observe the effect of facial affect on gaze behavior, facial affect had to be accurately produced by the RAs. To ensure the success of this modulation, video sequences of the 3 s of facial affect modulation were analyzed with the Facial Action Coding System (FACS; Ekman and Friesen, 1978) implemented in OpenFace (Amos et al., 2016). FACS is a comprehensive, anatomically based system for describing observable facial movements. It identifies specific facial muscle movements, known as Action Units (AUs), which correspond to particular emotions. The analysis focused on the mean intensities of three specific AUs:

- AU04 (Brow Lowerer): this action unit involves the lowering of the eyebrows, typically associated with negative emotions such as anger or sadness.
- AU06 (Cheek Raiser): commonly involved in the expression of happiness, contributing to the formation of crow's feet around the eyes.
- AU12 (Lip Corner Puller): also associated with happiness, resulting in the upward movement of the lip corners.

It was expected that positive affect would be dominated by activation of AU06 and AU12, whereas negative affect would be dominated by activation of AU04. In addition, the average time course of AU04 and AU06 activation over all participants was visualized to check for temporal deviations in affect modulation by the RAs. Thereby it was expected that activation of AUs would be highest during affect modulation and that there should be no pronounced temporal fluctuations in the onset of activation. After analysis of AUs, three trials of negative affect (2.21%) had to be excluded from the eye-tracking data analysis because the corresponding AUs were not activated.

### 2.5.4 Definition of areas of interest

All analyses of gaze data required the definition of areas of interest (AOIs) on the face of the RA to make assumption about the participant's gaze direction. We thus applied the facial recognition software OpenFace (Amos et al., 2016) and used the *Limited-Radius Voronoi Tessellation method* (LRVT; Hessels et al., 2016) to automatically generate these AOIs. Based on the data quality obtained in the study, we chose a radius of 1° visual angle (Vehlen et al., 2022) (see Figures 2C, E). We also added an ellipse around the face as an additional AOI (see Figure 2A). All available gaze data that were not directed to one of the facial AOIs were classified as AOI surrounding. Gaze data not available in the dataset were classified as missing. Missing data could be due to the participant looking outside the tracking area, blinking, or the eye-tracker losing the participant's eyes.

## 2.6 Statistical analyses

For each participant, all statistical analyses were based on eight trials of facial affect modulation (4× positive and 4× negative) within the social interaction, each lasting 9 s. Three seconds before the facial affect modulation, 3 s during and 3 s after (see Figure 1A). The length of the last segment was based on the minimum length of participants' responses to the questions in the social interaction.

We first analyzed gaze data based on aggregation using repeated-measures ANOVA. AOIs (eyes, nose, mouth, rest of face, surrounding), time (0–3,000 ms, 3,000–6,000 ms and 6,000–9,000 ms) and affect type (negative and positive) served as within-subject factors. In a second step, the AOI face (sum of AOI eyes, nose, mouth and rest of face) and the AOIs eyes and mouth were analyzed separately by means of repeated-measures ANOVA with time and affect type as within-subject factors. These analyses were conducted in R using the *afex* package (Singmann et al.,

2016) with the level of significance set to $p < 0.05$. In case the assumption of sphericity was violated, Greenhouse-Geisser correction was applied. Effect sizes are reported as partial eta squared ($\eta_p^2$). Planned comparisons with Bonferroni corrections were used to test whether differences in total dwell times on the face, eyes, and mouth actually occurred as a response to the affect modulation. $\eta_p^2$ values were estimated from the $t$-ratios as a measure of effect size.

The analyses were repeated with a focus of underlying dynamics in gaze data. For that, the total dwell times served as the basis for applying the *Smoothed Proportion of Looks Over Time* (SPLOT; Belopolsky, 2020) method. This method was developed as an alternative to classical AOI-based analysis, where temporal information is usually lost due to aggregation. The issue of multiple testing in these type of analyses is avoided by the application of cluster based permutation testing (Maris and Oostenveld, 2007; van Leeuwen et al., 2019). We implemented the method in Matlab (2018, version 9.4.0) and used it specifically to compare look probability between trials with negative and positive affect for the AOI face as well as the AOI eyes and mouth. For each comparison, we computed 1,000 iterations of permutation. During each iteration, the labels for the negative and positive affect trials were randomly permuted to break any association between the affect conditions and the observed gaze data. We then calculated the paired $t$-statistic for each permuted data set. A cluster was defined when two consecutive time points reached significance. We then used the summed $t$-statistics of the largest cluster to generate a null distribution of the test statistics. After creating the null distribution, we set a threshold for the clusters based on the 95th percentile of the distribution to identify significant differences. We than compared the two affect conditions by means of paired $t$-tests and again defined clusters. These clusters were considered significant only when the summed $t$-statistics were above the defined threshold derived from the distribution. The degree of smoothing was set to a standard deviation of 100 ms to correspond to the minimum duration of fixations (Belopolsky, 2020). We chose to use the average Cohen's $d$ for each cluster as a measure of effect size (Meyer et al., 2021). The code used for these analyses can be downloaded here: https://osf.io/x92ds/.

# 3 Results

## 3.1 Robustness, precision, accuracy

Gaze data quality analysis showed stable and sufficiently good quality over the course of the testing sessions (see Table 1). More particularly, the high accuracy allowed the discrimination between the gaze directed to the different facial features during affect modulation.

## 3.2 Validation of facial affect modulation

Mean intensity of AUs 04, 06 and 12 were used to validate the facial affect modulation—see Figure 1B. As expected, differences in AU04 activation emerged between affect types [$M \pm SD$; Negative: $1.27 \pm 0.35$; Positive: $0.07 \pm 0.13$; $F_{(1,34)} = 344.71$, $p < 0.001$,
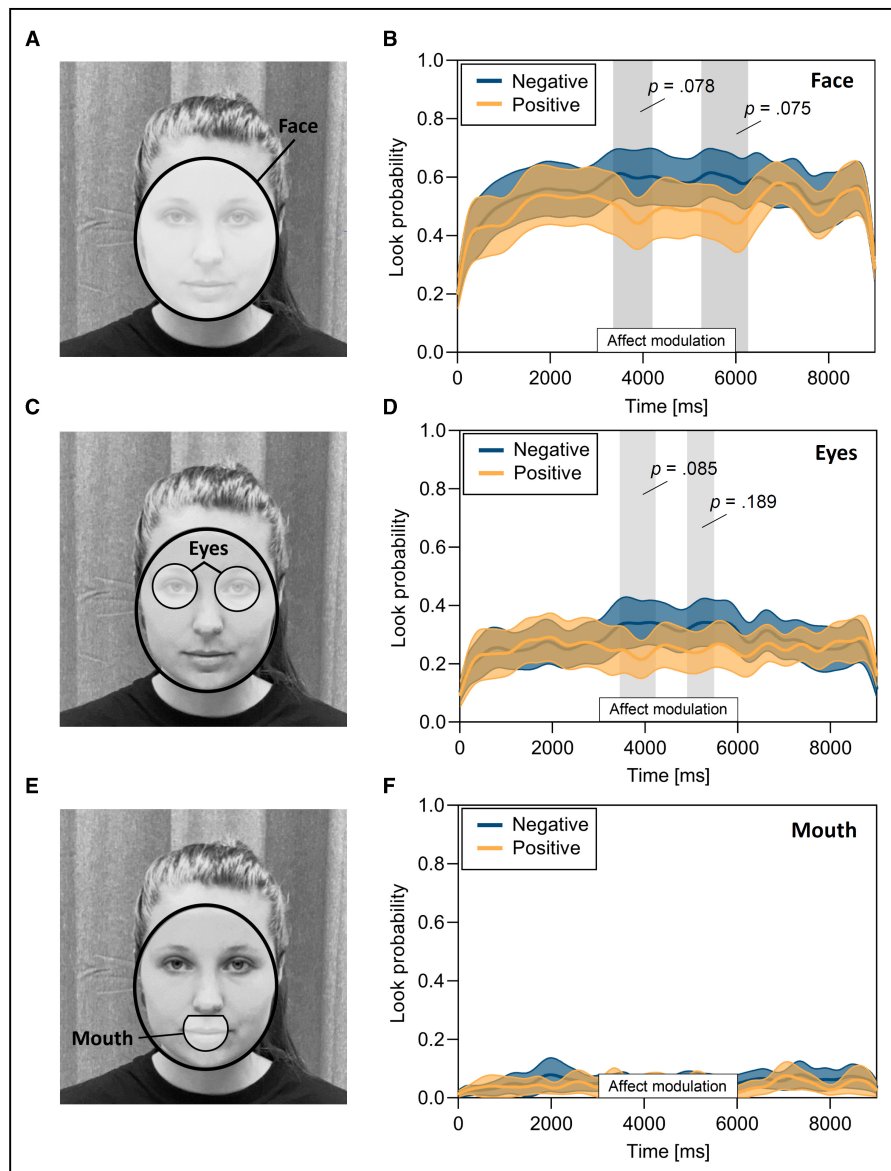
**FIGURE 2**
Look probability for facial areas of interest (AOIs) in response to positive and negative affect. **(A)** An ellipse around the face generated through landmark detection (AOI face) (Amos et al., 2016). **(B)** Look probability over time on the AOI face in response to facial affect modulation. **(C)** AOI eyes automatically generated with the Limited-Radius Voronoi Tessellation method (LRVT; Hessels et al., 2016) and a radius of 1° visual angle. **(D)** Look probability over time on the AOI eyes in response to facial affect modulation. **(E)** Automatically generated AOI mouth. **(F)** Look probability over time on the AOI mouth in response to facial affect modulation. Vertical gray shaded areas mark differences between conditions. Solid lines represent the mean look probability with the corresponding 95% confidence intervals. Written informed consent was obtained from the individual for the publication of the identifiable images.

$\eta_p^2 = 0.91$]. The same was true for activation of AU06 [$M \pm SD$; Negative: 0.43 ± 0.46; Positive: 1.40 ± 0.20; $F_{(1,34)} = 145.52$, $p < 0.001$, $\eta_p^2 = 0.81$], as well as AU12 [$M \pm SD$; Negative: 0.29 ± 0.37; Positive: 2.30 ± 0.38; $F_{(1,34)} = 834.07$, $p < 0.001$, $\eta_p^2 = 0.96$]. Furthermore, the time course analysis of AU activation during the 9 s of analysis (Figure 1A) confirmed that peak activation occurred during the 3 s of affect modulation, with moderate variance in the onset timing of these expressions (see Figure 1C). This indicates that the RAs consistently produced the intended facial affect modulation during the required period.

## 3.3 Aggregated measures of gaze behavior

The amount of missing gaze data (due to blinks, excessive head movements, viewing outside the tracking area, etc.) differed between affect types, with more missing values for positive compared to negative affect ($M \pm SD$; Positive: 38.20% ± 29.00%; Negative: 32.70% ± 27.10%; $t_{(410)} = 3.31$, $p = 0.001$, $r = 0.16$). All subsequent analyses were reported as total dwell times for specific AOIs relative to the total amount of valid data, i.e., total trial length minus missing data.

TABLE 1  Eye-tracking robustness, precision and accuracy determined within three validation sequences averaged over all target locations (wall, pre- & post-interaction).

|  | Wall | | Face (pre) | | Face (post) | |
|---|---|---|---|---|---|---|
|  | Mean | SD | Mean | SD | Mean | SD |
| **Robustness** | | | | | | |
| % valid data | 98.25 | 3.56 | 95.53 | 7.16 | 93.26 | 12.08 |
| **Precision (SD)** | | | | | | |
| In degrees | 0.32 | 0.09 | 0.44 | 0.17 | 0.44 | 0.18 |
| In cm | 0.73 | 0.21 | 1.01 | 0.39 | 1.01 | 0.41 |
| **Accuracy** | | | | | | |
| In degrees | 0.38 | 0.14 | 0.52 | 0.20 | 0.52 | 0.26 |
| In cm | 0.87 | 0.32 | 1.19 | 0.46 | 1.19 | 0.59 |

The data of quality indices published here have already been used for the evaluation of the used eye-tracking setup and can be found here: Vehlen et al. (2021) (Study 3).
Values in percent, degrees of visual angle and cm. Lower values reflect better precision and accuracy.

The first repeated-measures ANOVA (AOI × time × affect type) revealed only a significant main effect of AOIs, $F_{(3.14, 106.92)} = 19.29$, $\varepsilon = 0.79$, $p < 0.001$, $\eta_p^2 = 0.36$, with the eyes attracting the most attention on the face of the interaction partner ($M \pm SD$; Eyes = $37.90 \pm 28.40\%$ vs. nose = $12.2 \pm 18.20\%$, mouth = $6.81 \pm 13.70\%$ and rest of face = $17.70 \pm 19.20\%$). The main effect of time, $F_{(1.73, 58.66)} = 0.63$, $\varepsilon = 0.86$, $p = 0.514$, $\eta_p^2 = 0.02$ and affect type, $F_{(1, 34)} = 0.66$, $p = 0.422$, $\eta_p^2 = 0.02$ did not reach significance. Nor did any interaction effect [AOI × affect type: $F_{(1.93, 65.61)} = 1.94$, $\varepsilon = 0.48$, $p = 0.153$, $\eta_p^2 = 0.05$; AOI × time: $F_{(3.65, 124.23)} = 1.77$, $\varepsilon = 0.46$, $p = 0.146$, $\eta_p^2 = 0.05$; time × affect type: $F_{(1.98, 67.42)} = 1.20$, $p = 0.308$, $\eta_p^2 = 0.03$ and AOI × time × affect type: $F_{(3.78, 128.44)} = 1.37$, $\varepsilon = 0.47$, $p = 0.250$, $\eta_p^2 = 0.04$].

The following repeated-measures ANOVA for the AOI face showed no significant main effect of affect type, $F_{(1, 34)} = 3.28$, $p = 0.079$, $\eta_p^2 = 0.09$. Additionally, no main effect of time, $F_{(1.86, 63.41)} = 2.89$, $\varepsilon = 0.93$, $p = 0.067$, $\eta_p^2 = 0.08$ and no time × affect type interaction, $F_{(1.93, 65.59)} = 2.08$, $\varepsilon = 0.96$, $p = 0.135$, $\eta_p^2 = 0.06$ could be found. Planned comparisons revealed a significant difference in attention to the face between positive and negative affect during the affect modulation ($M \pm SD$; Negative: $79.20 \pm 28.31\%$; Positive: $70.17 \pm 32.63\%$; $p = 0.021$, $\eta_p^2 = 0.15$), but not before and after (all $p > 0.05$; see Figure 1D).

The same pattern emerged for the AOI eyes. The repeated-measures ANOVA revealed no significant main effect of affect type, $F_{(1, 34)} = 1.48$, $p = 0.232$, $\eta_p^2 = 0.05$. Again, no significant main effect of time, $F_{(1.77, 60.26)} = 0.87$, $\varepsilon = 0.89$, $p = 0.413$, $\eta_p^2 = 0.03$ as well as no significant time × affect type interaction were found, $F_{(1.75, 59.46)} = 2.96$, $\varepsilon = 0.87$, $p = 0.066$, $\eta_p^2 = 0.08$. As for the whole face, planned comparisons revealed a significant difference between negative and positive affect during the affect modulation ($M \pm SD$; Negative: $42.24 \pm 29.03\%$; Positive: $35.41 \pm 29.00\%$; $p = 0.024$, $\eta_p^2 = 0.14$) and again not for the time before and after modulation (all $p > 0.05$; see Figure 1E).

Furthermore, no significant main effect of affect type emerged for the AOI mouth, $F_{(1, 34)} = 3.05$, $p = 0.090$, $\eta_p^2 = 0.08$. On the

contrary, a significant main effect of time was found, $F_{(1.88, 63.76)} = 5.30$, $\varepsilon = 0.94$, $p = 0.009$, $\eta_p^2 = 0.14$. The interaction between affect type × time did not reach significance, $F_{(1.60, 54.27)} = 0.99$, $\varepsilon = 0.80$, $p = 0.363$, $\eta_p^2 = 0.03$. Planned comparisons revealed no differences in dwell times before, during and after positive compared to negative affect modulation (all $p > 0.05$; see Figure 1F).

## 3.4 Dynamic measures of gaze behavior

The SPLOT method was used to compare gaze behavior in response to negative and positive affect over the course of 9 s (3 s before, during and after affect modulation). Specifically, it was used to provide insight into temporal dynamics during the critical period of affect modulation. Thereby, two clusters emerged in the comparison between positive and negative affect for the AOI face and the AOI eyes, respectively. Differences in look probability on the face emerged around 3,358–4,208 ms (cluster $t$-statistic: 293.21, $p = 0.078$, $d = 0.48$) and 5,258–6,266 ms (cluster $t$-statistic: 305.45, $p = 0.075$, $d = 0.42$; see Figure 2B). Early differences in look probability on the AOI eyes occurred around the same time, namely 3,450–4,225 ms (cluster $t$-statistic: 251.13, $p = 0.085$, $d = 0.45$), and later around 4,908–5,492 (cluster $t$-statistic: 179.82, $p = 0.189$, $d = 0.43$; see Figure 2D). Differences between conditions for the look probability on the AOI mouth during the affect modulation did not occur (see Figure 2F). Although these differences did not reach statistical significance, they indicate attention-related trends in gaze behavior. Specifically, more attention appeared to be directed to the face, particularly the eyes, when the affect of the interaction partner was negative. Conversely, attention to these regions seemed to be reduced during positive affect. These trends align with our hypothesis that negative affect would result in increased attention to the eyes. However, contrary to our hypothesis, positive affect did not lead to greater attention to the mouth region.

## 4 Discussion

The study explored how facial affects influence gaze behavior in natural social interactions. Participants' gaze was recorded using remote eye-tracking during interactions with a RA, and responses to trained affect modulations were observed. A novel method using cluster-based permutation testing captured gaze dynamics alongside aggregated measures.

Facial affect modulations were observed during natural social interactions, which offered fewer opportunities for standardization (Hadders-Algra, 2022; Valtakari et al., 2021). To address the lack of a priori, independent validation of facial expressions, RAs were trained beforehand, and responses were scripted to ensure interaction standardization. Sequences of facial affect modulation were validated using an independent software that provides information about the activation of certain AUs (OpenFace; Amos et al., 2016). As hoped for, positive affect activated AUs linked to happiness (AU06 and AU12), while negative affect primarily triggered an AU associated with anger (AU04). The affect modulation was not subject to strong temporal fluctuations. However, despite RA's precise instructions, we avoided categorizing facial expressions beyond positive and negative.

As expected, the analysis of the aggregated measures revealed a general preference for the eye region over other facial AOIs for all affect types and over time. This highlights the eye region's significance as the most socially salient part of the face in social interactions (Birmingham et al., 2009; Devue et al., 2012; Tatler et al., 2010). Differences in missing values, which can be attributed to noise, blinking, or looking outside the tracking area, were observed across different affect types. Distinguishing between these events could clarify whether positive affect prompts more gaze aversion in this interaction, but an accurate video-based eye-tracking method is yet to be identified (Wisiecka et al., 2022).

In aggregated analyses, subtle differences emerged for the face, specifically the eyes, during negative affect modulation, aligning with increased eye gaze in standardized setups (Calvo et al., 2018). No significant attentional differences occurred for the AOI mouth, contradicting past studies on positive affect (Beaudry et al., 2014; Calvo et al., 2018; Green et al., 2003). However, analyzing only aggregate measurements can result in a loss of information within the high-frequency gaze signal (Belopolsky, 2020).

The SPLOT method maintains temporal resolution without raising the risk of false positives. While not achieving statistical significance, intriguing patterns emerged in participants' gaze behavior during affect modulation. Clusters of rapid changes in both positive and negative affect were identified likely associated with the first saccade in response to affect modulation. Our hypothesis that negative affect would increase attention to the eyes was partially supported by these patterns. Negative affect appeared to direct the gaze toward the face, specifically the eye region. However, contrary to our hypothesis, positive affect did not lead to greater attention to the mouth region, and instead, gaze appeared to be directed away from the face. A similar pattern at the end of the interaction suggested delayed disengagement for negative affect (Belopolsky et al., 2011). Findings regarding negative affect are consistent with those of a more standardized setup using dynamic stimuli (Calvo et al., 2018). Here, angry as well as sad stimuli drew more attention to the eye region. Simultaneously, this change in attention allocation contradicts results of studies using static stimuli (Beaudry et al., 2014; Hunnius et al., 2011). One explanation for these partially consistent results could be the importance of dynamic aspects in the study of facial affect, which is neglected in static stimuli (Krumhuber et al., 2013). Notably, heightened attention to negative affect could stem from its unpredictability in a positive interaction, aligning with research on stimulus unpredictability effects (Becker et al., 2007; Król and Król, 2017). In this context, positive affect could have been subtly noted, possibly explaining the contradictory gaze response. Moreover, participants, engaged in speaking during affect modulation and averted gaze could have enhanced cognitive resources for the heightened demands of communication (cognitive load theory; Doherty-Sneddon and Phelps, 2005; Glenberg et al., 1998).

Overall, however, the differences determined using the SPLOT method did not reach significance. Accordingly, the differences in gaze behavior in response to affect modulation were either too small or too brief. This again calls into question the validity of standardized setups in predicting natural gaze behavior (Hietanen et al., 2020; Laidlaw et al., 2011; Risko et al., 2012, 2016). The effects of the facial affect on gaze behavior from more standardized

setups could not be clearly replicated in our interactive eye-tracking study. This is in line with a recent study showing no effect of facial affect in a similar interactive setup (Pasqualette and Kulke, 2023). Besides the interaction potential, the influence of task on gaze behavior should be considered when interpreting the results (Hadders-Algra, 2022). While previous studies with a standardized experimental setup often presented the participants with a clear task (rate emotion, arousal, gaze direction, etc.), it can be assumed that natural interactions present evolving demands, which might effect gaze behavior. Additionally, the intensity and length of the gaze response to negative affect could also be attributed to the task, that was developed to generate interpersonal closeness between interaction partners (Aron et al., 1997) and during which negative affect might have been overlooked. However, questions included in the task (e.g., "If you could change one thing about the way you grew up, what would it be? Why?") can stimulate controversial discussions and potentially results in a negative affective response. Contextual factors may also have influenced the interpretation of emotional facial expressions. Although the questions were presented in a fixed order to maintain consistency, the content of the questions could interact with the displayed facial affect in ways that were not controlled for. For example, an angry facial expression may be perceived differently depending on whether the participant is discussing a neutral topic or a personal grievance. Although the presentation of facial affect was randomized to mitigate order effects, future studies should consider further controlling the context in which emotional expressions are presented.

Another explanation for the absence of significant differences could be the reduced statistical power resulting from a relatively small sample size. Determining the appropriate sample size was challenging because of the considerable variability in effect sizes reported for the interaction effect in previous studies (e.g. AOI × affect: $\eta_p^2 = 0.09$; Eisenbarth and Alpers, 2011, or $\eta_p^2 = 0.48$; Calvo et al., 2018) and the lack of meta-analyses. In addition, it was difficult to estimate the influence of the naturalistic setting on the anticipated effects. Furthermore, we wanted to look at the data dynamically; to our knowledge, no previous studies have observed the effects of affect over time. This also made it difficult to predict the exact sample size, as gaze behavior is expected to become significantly noisier over time. Combined with a complex study design that limited the feasible sample size, we decided to address threats to statistical power through increasing measurement precision (Hedge et al., 2018; Nebe et al., 2023; Smith and Little, 2018). This involved excluding participants with compromised gaze data quality, using a validated eye-tracking setup (Tönsing et al., 2022; Vehlen et al., 2021), assessing data quality twice (pre- and post-interaction), and validating facial affect modulation.

Apart from the above limitations, the present data, together with several other recent studies, suggest potential differences in gaze behavior during social interactions compared with gaze behavior in more standardized setups. Brief facial affect modulation within positive social interactions appears to have short-term rather than fundamental effects on healthy participants' gaze behavior. To validate these preliminary findings and enhance their generalizability, subsequent studies should replicate this

research with larger sample sizes. Moreover, future research should explore gaze behavior across diverse social interactions, encompassing a broader range of facial expressions. Furthermore, the additional analysis of the interaction partners' gaze behavior can illuminate the interplay of these signals in the interactional context. Together with information from other modalities, such as the analysis of speech, this can ultimately enable a deeper understanding of dynamic systems such as social interactions. Consequently, these findings offer valuable initial insights into the impact of affect on gaze behavior, providing a foundation for future inquiries into effect sizes, analytical approaches, and measurement precision.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/x92ds/.

## Ethics statement

The studies involving humans were approved by Ethics Committee of the University of Trier. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

AV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. AB: Software, Validation, Writing – original draft, Writing – review & editing. GD: Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1433483/full#supplementary-material

## References

Amos, B., Ludwiczuk, B., and Satyanarayanan, M. (2016). *OpenFace: A general-purpose face recognition library with mobile applications*. CMU-CS-16-118. Pittsburgh, PA: CMU School of Computer Science.

Argyle, M., and Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.

Aron, A., Melinat, E., Aron, E. N., Vallon, R. D., and Bator, R. J. (1997). The experimental generation of interpersonal closeness: a procedure and some preliminary findings. *Pers. Soc. Psychol. Bull.* 23, 363–377. doi: 10.1177/0146167297234003

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., and Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cogn. Emot.* 28, 416–432. doi: 10.1080/02699931.2013.833500

Becker, M. W., Pashler, H., and Lubin, J. (2007). Object-intrinsic oddities draw early saccades. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 20–30. doi: 10.1037/0096-1523.33.1.20

Belopolsky, A. V. (2020). "Getting more out of Area of Interest (AOI) analysis with SPLOT," in *ACM Symposium on Eye Tracking Research and Applications* (New York, NY: ACM), 1–4. doi: 10.1145/3379156.3391372

Belopolsky, A. V., Devue, C., and Theeuwes, J. (2011). Angry faces hold the eyes. *Vis. Cogn.* 19, 27–36. doi: 10.1080/13506285.2010.536186

Birmingham, E., Bischof, W. F., and Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vis. Res.* 49, 2992–3000. doi: 10.1016/j.visres.2009.09.014

Calvo, M. G., Fernández-Martín, A., Gutiérrez-García, A., and Lundqvist, D. (2018). Selective eye fixations on diagnostic face regions of dynamic emotional expressions: KDEF-dyn database. *Sci. Rep.* 8:17039. doi: 10.1038/s41598-018-35259-w

Carter, B. T., and Luke, S. G. (2020). Best practices in eye tracking research. *Int. J. Psychophysiol.* 155, 49–62. doi: 10.1016/j.ijpsycho.2020.05.010

Chen, J., van den Bos, E., and Westenberg, P. M. (2020). A systematic review of visual avoidance of faces in socially anxious individuals: influence of severity, type of social situation, and development. *J. Anxiety Disord.* 70:102193. doi: 10.1016/j.janxdis.2020.102193

Dalmaso, M., Castelli, L., and Galfano, G. (2020). Social modulators of gaze-mediated orienting of attention: a review. *Psychon. Bull. Rev.* 27, 833–855. doi: 10.3758/s13423-020-01730-x

De Lillo, M., Foley, R., Fysh, M. C., Stimson, A., Bradford, E. E. F., Woodrow-Hill, C., et al. (2021). Tracking developmental differences in real-world social attention across adolescence, young adulthood and older adulthood. *Nat. Hum. Behav.* 5, 1381–1390. doi: 10.1038/s41562-021-01113-9

Devue, C., Belopolsky, A. V., and Theeuwes, J. (2012). Oculomotor guidance and capture by irrelevant faces. *PLOS ONE* 7:e34598. doi: 10.1371/journal.pone.0034598

Doherty-Sneddon, G., and Phelps, F. G. (2005). Gaze aversion: a response to cognitive or social difficulty? *Mem. Cogn.* 33, 727–733. doi: 10.3758/BF03195338

Eisenbarth, H., and Alpers, G. W. (2011). Happy mouth and sad eyes: scanning emotional facial expressions. *Emotion* 11, 860–865. doi: 10.1037/a0022758

Ekman, P., and Friesen, W. V. (1978). *Facial Action Coding System.* Palo Alto, CA: Consulting Psychologists Press.

Glenberg, A. M., Schroeder, J. L., and Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Mem. Cogn.* 26, 651–658. doi: 10.3758/BF03211385

Green, M. J., Williams, L. M., and Davidson, D. (2003). In the face of danger: specific viewing strategies for facial expressions of threat? *Cogn. Emot.* 17, 779–786. doi: 10.1080/02699930302282

Grossman, R. B., Zane, E., Mertens, J., and Mitchell, T. (2019). Facetime vs. screentime: gaze patterns to live and video social stimuli in adolescents with ASD. *Sci. Rep.* 9:12643. doi: 10.1038/s41598-019-49039-7

Hadders-Algra, M. (2022). Human face and gaze perception is highly context specific and involves bottom-up and top-down neural processing. *Neurosci. Biobehav. Rev.* 132, 304–323. doi: 10.1016/j.neubiorev.2021.11.042

Hedge, C., Powell, G., and Sumner, P. (2018). The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* 50, 1166–1186. doi: 10.3758/s13428-017-0935-1

Hessels, R. S. (2020). How does gaze to faces support face-to-face interaction? A review and perspective. *Psychon. Bull. Rev.* 27, 856–881. doi: 10.3758/s13423-020-01715-w

Hessels, R. S., Cornelissen, T. H. W., Hooge, I. T. C., and Kemner, C. (2017). Gaze behavior to faces during dyadic interaction. *Can. J. Exp. Psychol.* 71, 226–242. doi: 10.1037/cep0000113

Hessels, R. S., Kemner, C., van den Boomen, C., and Hooge, I. T. C. (2016). The area-of-interest problem in eyetracking research: a noise-robust solution for face and sparse stimuli. *Behav. Res. Methods* 48, 1694–1712. doi: 10.3758/s13428-015-0676-y

Hietanen, J. O., Peltola, M. J., and Hietanen, J. K. (2020). Psychophysiological responses to eye contact in a live interaction and in video call. *Psychophysiology* 57:e13587. doi: 10.1111/psyp.13587

Horn, A., Mergenthaler, L., and Gamer, M. (2022). Situational and personality determinants of social attention in a waiting room scenario. *Vis. Cogn.* 30, 86–99. doi: 10.1080/13506285.2021.1966151

Hunnius, S., de Wit, T. C. J., Vrins, S., and von Hofsten, C. (2011). Facing threat: infants' and adults' visual scanning of faces with neutral, happy, sad, angry, and fearful emotional expressions. *Cogn. Emot.* 25, 193–205. doi: 10.1080/15298861003771189

Itier, R. J., and Batty, M. (2009). Neural bases of eye and gaze processing: the core of social cognition. *Neurosci. Biobehav. Rev.* 33, 843–863. doi: 10.1016/j.neubiorev.2009.02.004

Just, M. A., and Carpenter, P. A. (1984). "Using eye fixations to study reading comprehension," in *New Methods in Reading Comprehension Research*, eds. D. E. Kieras, and M. A. Just (Mahwah, NJ: Erlbaum), 151–182.

Król, M. E., and Król, M. (2017). The trickle-down effect of predictability: secondary task performance benefits from predictability in the primary task. *PLOS ONE* 12:e0180573. doi: 10.1371/journal.pone.0180573

Krumhuber, E. G., Kappas, A., and Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: a review. *Emot. Rev.* 5, 41–46. doi: 10.1177/1754073912451349

Kulke, L., Ertugrul, S., Reyentanz, E., and Thomas, V. (2023). Uncomfortable staring? Gaze to other people in social situations is inhibited in both infants and adults. *Dev. Sci.* 27:e13468. doi: 10.1111/desc.13468

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., and Kingstone, A. (2011). Potential social interactions are important to social attention. *Proc. Natl. Acad. Sci.* 108, 5548–5553. doi: 10.1073/pnas.1017022108

Loughland, C. M., Williams, L. M., and Gordon, E. (2002). Schizophrenia and affective disorder show different visual scanning behavior for faces: a trait versus state-based distinction? *Biol. Psychiatry* 52, 338–348. doi: 10.1016/S0006-3223(02)01356-2

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024

Meyer, M., Lamers, D., Kayhan, E., Hunnius, S., and Oostenveld, R. (2021). Enhancing reproducibility in developmental EEG research: BIDS, cluster-based permutation tests, and effect sizes. *Dev. Cogn. Neurosci.* 52:101036. doi: 10.1016/j.dcn.2021.101036

Nebe, S., Reutter, M., Baker, D., Bölte, J., Domes, G., Gamer, M., et al. (2023). Enhancing precision in human neuroscience. *PsyArXiv.* doi: 10.31234/osf.io/m8c4k

Pasqualette, L., and Kulke, L. (2023). Effects of emotional content on social inhibition of gaze in live social and non-social situations. *Sci. Rep.* 13:1. doi: 10.1038/s41598-023-41154-w

Risko, E. F., Laidlaw, K. E., Freeth, M., Foulsham, T., and Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Front. Hum. Neurosci.* 6:143. doi: 10.3389/fnhum.2012.00143

Risko, E. F., Richardson, D. C., and Kingstone, A. (2016). Breaking the fourth wall of cognitive science: real-world social attention and the dual function of gaze. *Curr. Dir. Psychol.* 25, 70–74. doi: 10.1177/0963721415617806

Rösler, L., Göhring, S., Strunz, M., and Gamer, M. (2021). Social anxiety is associated with heart rate but not gaze behavior in a real social interaction. *J. Behav. Ther. Exp. Psychiatry* 70:101600. doi: 10.1016/j.jbtep.2020.101600

Senju, A., and Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn. Sci.* 13, 127–134. doi: 10.1016/j.tics.2008.11.009

Senju, A., Vernetti, A., Kikuchi, Y., Akechi, H., Hasegawa, T., and Johnson, M. H. (2013). Cultural background modulates how we look at other persons' gaze. *Int. J. Behav. Dev.* 37, 131–136. doi: 10.1177/0165025412465360

Setien-Ramos, I., Lugo-Marín, J., Gisbert-Gustemps, L., Díez-Villoria, E., Magán-Maganto, M., Canal-Bedia, R., et al. (2022). Eye-tracking studies in adults with autism spectrum disorder: a systematic review and meta-analysis. *J. Autism Dev. Disord.* 53, 2430–2443. doi: 10.1007/s10803-022-05524-z

Singmann, H., Bolker, B., Westfall, J., and Aust, F. (2016). *afex: Analysis of Factorial Experiments*. R package version 1.4.1. Available at: https://CRAN.R-project.org/package=afex

Smith, P. L., and Little, D. R. (2018). Small is beautiful: in defense of the small-N design. *Psychon. Bul. Rev.* 25, 2083–2101. doi: 10.3758/s13423-018-1451-8

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., and Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *I-Perception* 1, 7–27. doi: 10.1068/i0382

Tönsing, D., Schiller, B., Vehlen, A., Spenthof, I., Domes, G., and Heinrichs, M. (2022). No evidence that gaze anxiety predicts gaze avoidance behavior during face-to-face social interaction. *Sci. Rep.* 12:1. doi: 10.1038/s41598-022-25189-z

Valtakari, N. V., Hooge, I. T. C., Viktorsson, C., Nyström, P., Falck-Ytter, T., and Hessels, R. S. (2021). Eye tracking in human interaction: possibilities and limitations. *Behav. Res. Methods* 53, 1592–1608. doi: 10.3758/s13428-020-01517-x

van Leeuwen, J., Smeets, J. B. J., and Belopolsky, A. V. (2019). Forget binning and get SMART: getting more out of the time-course of response data. *Atten. Percept. Psychophys.* 81, 2956–2967. doi: 10.3758/s13414-019-01788-3

Vehlen, A., Spenthof, I., Tönsing, D., Heinrichs, M., and Domes, G. (2021). Evaluation of an eye tracking setup for studying visual attention in face-to-face conversations. *Sci. Rep.* 11:1. doi: 10.1038/s41598-021-81987-x

Vehlen, A., Standard, W., and Domes, G. (2022). How to choose the size of facial areas of interest in interactive eye tracking. *PLOS ONE* 17:e0263594. doi: 10.1371/journal.pone.0263594

Wisiecka, K., Krejtz, K., Krejtz, I., Sromek, D., Cellary, A., Lewandowska, B., et al. (2022). "Comparison of webcam and remote eye tracking [Paper presentation]," in *ETRA '22: 2022 Symposium on Eye Tracking Research and Applications* (Seattle, WA). doi: 10.1145/3517031.3529615

Wright, R. D., and Ward, L. M. (2008). *Orienting of Attention.* Oxford: Oxford University Press.

Yarbus, A. L. (1967). *Eye Movements and Vision.* New York, NY: Springer.