
Eye-Tracking During Computerized Social Cognition Tasks in Healthy Adults: Feasibility and Task-Dependent Gaze Patterns in a Pilot Study

[Rosa Ayesa-Arriola](#) , [Manuel Sevilla-Ramos](#) , Esther Setién-Suero , [Luis Rodríguez-Cobo](#) , [Susana Ochoa-Rodríguez](#) , [Alexandre Díaz-Pons](#) *

Posted Date: 18 May 2026

doi: 10.20944/preprints202605.1098.v1

Keywords: eye-tracking technology; fixation; ocular; eye movements; social cognition; facial recognition; facial expression; neuropsychological tests; psychometrics; pilot projects



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Eye-Tracking During Computerized Social Cognition Tasks in Healthy Adults: Feasibility and Task-Dependent Gaze Patterns in a Pilot Study

Rosa Ayesa-Arriola ^{1,2}, Manuel Sevilla-Ramos ³, Esther Setién-Suero ¹, Luis Rodríguez-Cobo ^{4,5,6}, Susana Ochoa-Rodríguez ^{2,7,8} and Alexandre Díaz-Pons ^{1,9,*}

¹ Grupo de Investigación en Enfermedades Mentales, Instituto de Investigación Sanitaria Valdecilla (IDIVAL), 39011 Santander, España

² Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Instituto de Salud Carlos III, 28029 Madrid, España

³ Facultad de Psicología, Universidad de Salamanca (USAL), 37005 Salamanca, España

⁴ Grupo de Investigación en Ingeniería Fotónica, Instituto de Investigación Sanitaria Valdecilla (IDIVAL), 39011 Santander, España

⁵ Grupo de Ingeniería Fotónica (GIF), Universidad de Cantabria (UC), 39005 Santander, España

⁶ Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, 28029 Madrid, España

⁷ Parc Sanitari Sant Joan de Déu, 08830 Sant Boi de Llobregat, Barcelona, España

⁸ Grup MERITT, Institut de Recerca Sant Joan de Déu, 08950 Esplugues de Llobregat, Barcelona, España

⁹ Escuela de Doctorado de la Universidad de Cantabria (EDUC), Universidad de Cantabria (UC), 39005 Santander, España

* Correspondence: alexandre.diaz@idival.org

Abstract

Background: Social-cognition assessment often relies on endpoint measures such as accuracy, which provide limited information about how social stimuli are visually sampled. Eye-tracking can capture visual-sampling processes, but the meaning of gaze metrics depends on task structure. **Objective:** To examine the feasibility and preliminary informativeness of eye-tracking during two computerized social-cognition tasks in healthy adults. **Methods:** Nineteen healthy adults completed a full-face facial emotion-recognition task (TREC) and the Reading the Mind in the Eyes Test (RMET) while gaze was recorded. Measures included fixation count, cumulative fixation duration, and reaction time. TREC analyses examined gaze allocation across the eyes, nose, mouth, and facial hemifields. Analyses were exploratory and hypothesis-generating. **Results:** In the TREC, gaze was mainly allocated to the eyes and nose, with less sampling of the mouth. Higher TREC performance was accompanied by greater eye-region and left-hemiface viewing. Negative expressions elicited more fixations, and older participants showed greater eye-region sampling. In the RMET, participants showed higher fixation count, longer cumulative fixation duration, and longer response time than in the TREC, but gaze metrics were not clearly associated with demographic or performance variables. **Conclusions:** Eye-tracking was feasible and yielded coherent, task-dependent visual-sampling patterns in this small pilot sample. Full-face stimuli enabled spatially resolved gaze characterization, whereas eye-region stimuli mainly provided global inspection metrics. Findings are preliminary and should inform larger studies testing the clinical or mechanistic value of gaze-derived measures.

Keywords: eye-tracking technology; fixation; ocular; eye movements; social cognition; facial recognition; facial expression; neuropsychological tests; psychometrics; pilot projects

1. Introduction

Social cognition refers to the neurocognitive capacity to extract meaning from other agents by integrating perceptual, attentional, affective, and inferential processes. It comprises partially dissociable domains—including social perception, facial emotion recognition, affective processing, theory of mind, empathy, attributional style, and social knowledge—that allow individuals to detect socially relevant cues, evaluate their significance, and adjust behavior accordingly[1–3]. Faces and eyes are privileged social signals through which individuals infer emotion, intention, trust, threat, affiliation, and mental states[4,5]. Although human social cognition is expanded by language, symbolic representation, cultural learning, and explicit mental-state reasoning, it builds on conserved social-attentional mechanisms shared across social species, especially sensitivity to faces, gaze direction, emotional displays, and approach–avoidance cues[2,6–8]. Neurobiologically, these functions depend on distributed perceptual, salience-related, limbic, and associative cortical systems that support facial encoding, affective valuation, contextual integration, and social prediction[1,9]. Across development, these systems are progressively calibrated through experience[10], and cross-cultural evidence indicates that face-scanning strategies may vary in the relative use of eye, mouth, and central facial information[11,12]. Thus, social-cognitive performance is not only a matter of giving a correct or incorrect response; the same behavioral outcome may arise from different configurations of perceptual sensitivity, attentional allocation, affective salience, and inferential processing[2,13]. This makes social cognition clinically important and methodologically complex: to understand performance, it is necessary to examine not only what response is given, but how social information is sampled before that response emerges[13–15].

Despite this complexity, social-cognition assessment remains largely anchored in endpoint behavioral metrics, especially accuracy scores from facial emotion-recognition tasks and the Reading the Mind in the Eyes Test[16,17]. These instruments have been essential for documenting impairments across schizophrenia-spectrum disorders, autism, mood disorders, neurodegenerative diseases, and other conditions affecting social functioning[3,9,18,19]. However, accuracy scores provide limited mechanistic information. A correct or incorrect response compresses several processing stages—perceptual encoding, attentional selection, affective evaluation, evidence accumulation, and inferential decision-making—into a single outcome[2]. As a result, equivalent performance may conceal different strategies, and similar impairment may reflect disruptions at different levels of the social-cognitive system[3,13]. This limitation is especially relevant in heterogeneous or compensatory contexts. Some individuals may achieve preserved performance through slower, broader, or more effortful exploration; others may fail because they do not sample the most informative facial regions, misattribute affective salience, or struggle to integrate perceptual evidence into social meaning[20]. Task structure further complicates interpretation: full-face paradigms, eye-region tasks, static images, dynamic stimuli, and different response formats impose distinct perceptual and inferential constraints[16,17]. Therefore, performance cannot be treated as a transparent measure of social cognition without considering how the task organizes information sampling. This has motivated growing interest in process-level measures that can characterize how socially meaningful evidence is acquired before an explicit response is made[13].

Eye-tracking is well suited to address this gap because it captures visual sampling during task execution[21]. Fixation location, fixation duration, scan patterns, and response timing can reveal how attention is allocated across socially informative facial features and whether performance appears to rely on focused extraction of relevant cues, diffuse exploration, or compensatory effort[13]. Its value, however, depends on task design. Full-face emotion-recognition tasks provide a spatially rich field in which gaze can be distributed across the eyes, nose, mouth, and facial laterality; eye-region tasks, by contrast, restrict visual exploration while increasing the inferential demand placed on limited information[4,17]. The present pilot study examines the feasibility and preliminary informativeness of integrating eye-tracking into computerized social-cognition assessment in healthy adults, using a full-face facial emotion-recognition task and the Reading the Mind in the Eyes Test. Rather than establishing diagnostic thresholds or normative standards, it adopts a signal-characterization approach: determining whether gaze-derived metrics produce coherent, interpretable patterns under

controlled conditions and whether task format shapes the granularity of information obtained. Establishing such reference patterns in non-clinical participants is a necessary step before testing whether clinical deviations reflect altered perceptual encoding, attentional prioritization, affective salience, processing efficiency, or compensatory visual strategies[18,19].

2. Methods

Participants

The pilot sample was recruited by the IDIVAL Mental Illness Research Department's neuropsychology lab at the Marqués de Valdecilla Research Institute (IDIVAL), Santander, Spain, as part of the Spanish National Research Project PI18/00212, "Influence of attachment style on social cognition and cognitive biases in individuals with first-episode psychosis, chronic schizophrenia, and healthy controls." The final sample comprised 19 healthy adults (11 males and 8 females), all of whom received detailed information about the study and provided written informed consent. The study protocol was approved by CEIC-Cantabria (internal code: 2019.024). Inclusion criteria were:

- i) Absence of any mental disorder and related treatment during the previous year, assessed using the Composite International Diagnostic Interview (CIDI).
- ii) Age between 18 and 65 years.
- iii) Absence of traumatic brain injury, dementia, or intellectual disability (IQ < 70).
- iv) Absence of substance abuse.

Social cognition tasks

Participants completed two computerized social-cognition tasks: the Spanish version of the *Test de Reconocimiento Emocional en Caras* (TREC) and the Spanish adaptation of the *Reading the Mind in the Eyes Test* (RMET). The two tasks were selected because they impose different perceptual constraints on social-information sampling: the TREC presents full-face emotional stimuli, allowing gaze allocation to be examined across multiple facial regions, whereas the RMET restricts visual input to the eye region and therefore places greater inferential demand on limited perceptual information.

- i) **Facial Emotion Recognition Task (TREC)**[22,23]. The Spanish version of the TREC was administered as a computerized full-face facial emotion-recognition task. The task consists of 20 items, each corresponding to a photograph of the face of a Caucasian woman. Half of the items depict emotional expressions of negative valence and half depict emotional expressions of positive valence. For each item, participants were required to identify the emotion expressed by the face by selecting one of two response options. In the present study, the TREC was used to assess facial emotion recognition under spatially rich visual conditions, allowing eye-tracking analyses of region-specific gaze allocation across the eyes, nose, mouth, and facial hemifields.
- ii) **Reading the Mind in the Eyes Test (RMET)**[24,25]. The Spanish adaptation of the RMET was also administered in computerized format. The task consists of 36 black-and-white images originally taken from magazines, showing only the eye and eyebrow region of different Caucasian individuals, with an equal number of male and female faces. Each item depicts a complex mental or emotional state. Participants were required to select, from four response alternatives, the word that best described what the person in the image was thinking or feeling. In contrast to the TREC, the RMET restricts visual information to the eye region, preventing equivalent comparisons across full-face areas of interest and allowing assessment of global visual inspection during mental-state attribution under constrained perceptual input.

Eye-tracking measures

Gaze behavior was recorded during both computerized social-cognition tasks using a Tobii Pro Nano eye tracker integrated into custom software, enabling synchronized stimulus presentation, response registration, and gaze-data acquisition. Three eye-tracking variables were extracted:

- i) **Fixation count** was defined as the number of discrete fixations registered during a task or within a given area of interest, indexing how often visual attention was allocated to that region.
- ii) **Cumulative fixation duration** was defined as the total time, in milliseconds, spent fixating on a task or area of interest, indexing the amount of visual inspection devoted to that stimulus region.
- iii) **Reaction time** was defined as the time elapsed between stimulus onset and participant response, indexing response speed during social-cognitive judgment.

For the full-face emotion-recognition task, predefined areas of interest were manually delimited around the eyes, nose, and mouth to quantify region-specific visual sampling. In addition, each face was divided into left and right hemifacial regions to examine gaze laterality. For each area of interest, fixation count and cumulative fixation duration were extracted and compared across facial regions, emotional valence, sex, age, and task performance. For the RMET, gaze measures were computed at the task level because stimuli were restricted to the eye region, preventing equivalent comparisons across facial areas. This allowed assessment of global visual inspection during eye-region mental-state attribution while preserving the distinction between spatially rich and visually constrained task formats.

The computerized assessment was implemented within a multi-screen supervisor-participant architecture in which task presentation, response registration, gaze-data acquisition, and real-time monitoring could be synchronized during administration (Figure 1). In the present pilot study, the analytic focus was restricted to behavioral and eye-tracking-derived measures.

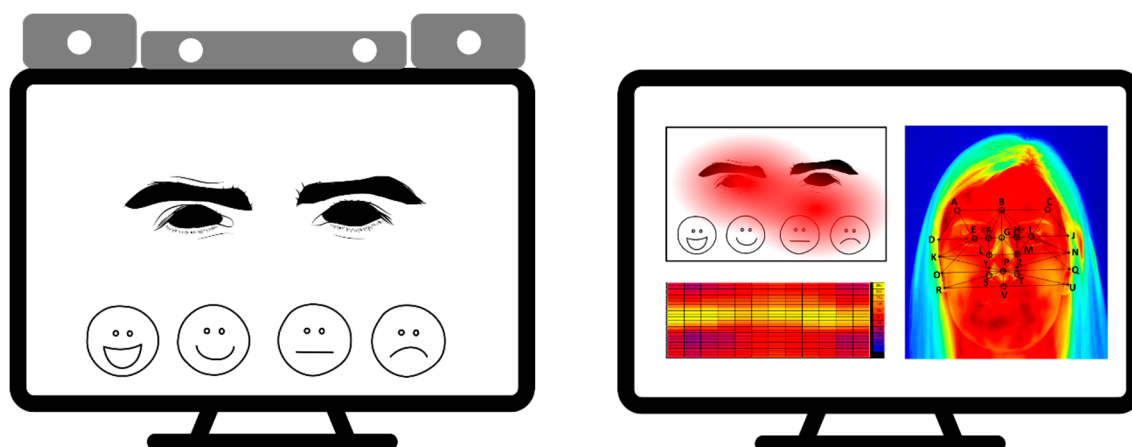


Figure 1. Proposed multi-screen architecture for computerized social-cognition assessment with real-time biometric monitoring. Schematic representation of the experimental setup. The participant-facing display presented the computerized social-cognition tasks and response options, while the supervisor interface-controlled task administration and enabled real-time monitoring of acquisition streams and task progression. The system was designed to synchronize stimulus presentation, response registration, gaze-data recording, and concurrent signal visualization, allowing subsequent extraction of behavioral and eye-tracking-derived measures, including fixation count, cumulative fixation duration, reaction time, and area-of-interest metrics. In the present pilot study, analyses focused on gaze-derived and behavioral variables. The diagram is schematic and does not display participant-identifiable data.

Statistical analysis

Statistical analyses were performed using SPSS version 23.0. Descriptive statistics were used to summarize sociodemographic, behavioral, and eye-tracking variables. Continuous variables are reported as means and standard deviations, and categorical variables as frequencies and percentages. Within-task comparisons across facial areas of interest and emotional valence were conducted using paired-samples *t*-tests when the same participants contributed data to both conditions. Between-group comparisons were conducted using independent-samples *t*-tests for sex, age group, and task-

performance group comparisons. Pearson correlation coefficients were calculated to examine associations between behavioral performance, age, education, and gaze-derived measures. Effect sizes were estimated using Cohen's *d*. Given the pilot nature of the study and the small sample size, all analyses were interpreted as exploratory and hypothesis-generating. Statistical significance was set at $p < .05$.

3. Results

The sample included 19 healthy adults, of whom 11 were men and 8 were women. Descriptive characteristics are shown in Table 1. Mean age was 35.89 years ($SD = 11.35$), and mean education was 13.16 years ($SD = 2.32$). Mean TREC score was 18.00 ($SD = 1.49$), and mean RMET score was 25.10 ($SD = 2.20$). No significant sex differences were observed in age, years of education, TREC performance, or RMET performance (all $p > .05$). Most participants were right-handed (89.47%).

During the TREC, fixation count differed across facial areas of interest (Table 2; Figure 1B–C). In the total sample, participants showed a mean of 89.35 fixations on the eyes ($SD = 51.40$), 83.26 on the nose ($SD = 32.96$), and 21.43 on the mouth ($SD = 16.84$). Post hoc comparisons showed significantly more fixations on the eyes than on the mouth ($t = -17.58$, $p < .001$) and on the nose than on the mouth ($t = -16.00$, $p < .001$). The corresponding Cohen's *d* values were 0.14 for eyes versus nose, 1.78 for eyes versus mouth, and 2.36 for nose versus mouth.

Cumulative fixation duration also differed across TREC facial areas of interest. In the total sample, cumulative fixation duration was 210.63 ms for the eyes ($SD = 205.29$), 173.16 ms for the nose ($SD = 115.83$), and 49.60 ms for the mouth ($SD = 50.35$). Post hoc comparisons showed greater cumulative fixation duration on the eyes than on the mouth ($t = -13.94$, $p < .001$) and on the nose than on the mouth ($t = -10.70$, $p < .001$). The corresponding Cohen's *d* values were 0.22 for eyes versus nose, 1.08 for eyes versus mouth, and 1.38 for nose versus mouth.

TREC performance-group analyses showed differences in eye-region gaze measures. Participants scoring above the TREC mean showed more fixations on the eyes than participants scoring at or below the TREC mean ($t = -2.15$, $p = .05$, $d = 0.97$). They also showed greater cumulative fixation duration on the eyes ($t = -2.33$, $p = .03$, $d = 1.00$).

TREC gaze laterality by performance group is shown in Table 3 and Figure 1D. Participants scoring above the TREC mean showed more fixations on the left hemiface ($M = 154.47$, $SD = 64.63$) than participants scoring at or below the TREC mean ($M = 33.68$, $SD = 31.37$; $t = -5.42$, $p < .001$). Participants scoring at or below the TREC mean showed more fixations on the right hemiface ($M = 137.63$, $SD = 39.02$) than participants scoring above the TREC mean ($M = 70.84$, $SD = 56.14$; $t = 3.07$, $p = .007$). TREC accuracy was positively correlated with left-side facial fixations ($r = .64$, $p = .003$).

No significant sex differences were observed in the spatial pattern of TREC gaze allocation. Women showed shorter TREC response times than men at trend level (3.33 vs. 4.05 seconds; $t = -1.96$, $p = .06$, $d = 0.88$). Participants older than 35 years showed more fixations on the eyes ($t = -2.30$, $p = .03$, $d = 1.03$) and greater cumulative fixation duration on the eyes ($t = -2.12$, $p = .05$, $d = 0.91$). Age was positively correlated with fixation count on the eyes ($r = .63$, $p = .003$) and cumulative fixation duration on the eyes ($r = .49$, $p = .03$). Negative expressions elicited more fixations than positive expressions ($t = -2.65$, $p = .02$, $d = 0.56$). Cumulative fixation duration did not differ significantly by emotional valence.

In the RMET, participants showed a mean fixation count of 467.16 ($SD = 171.99$) and a mean cumulative fixation duration of 2215.87 ms ($SD = 1346.71$; Table 4). No significant differences in RMET gaze metrics were observed according to sex, age, education, performance level, or emotional valence. Participants with higher educational level showed higher RMET scores at trend level ($t = -2.10$, $p = .05$, $d = 0.99$).

At the task level, the RMET showed higher fixation count, longer cumulative fixation duration, and longer response time than the TREC (Figure 1A). Mean response time was 7.50 seconds for the RMET and 3.75 seconds for the TREC.

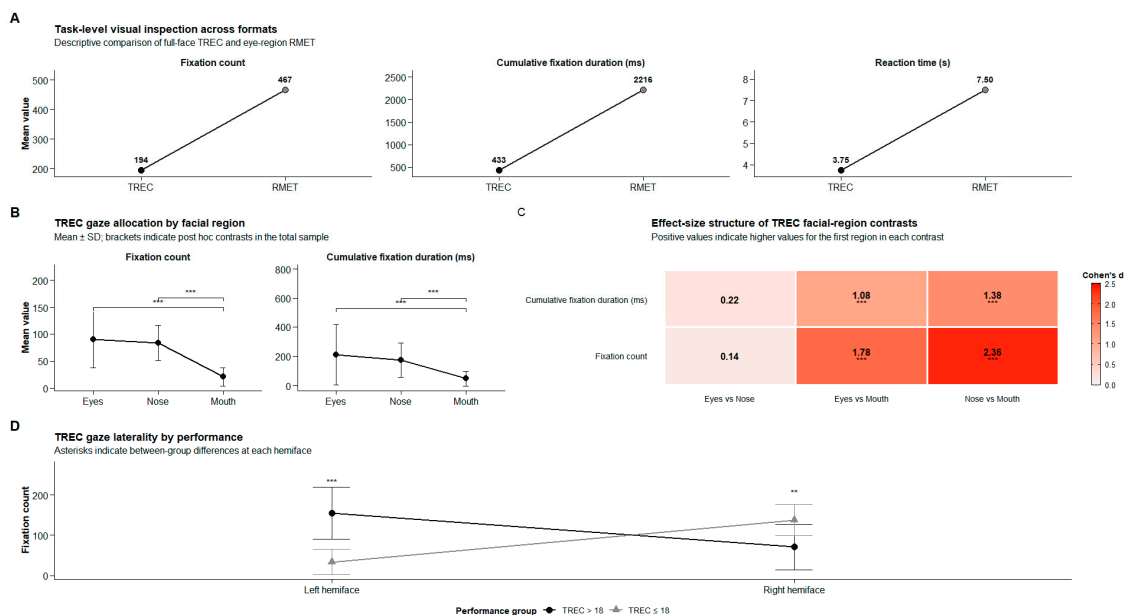


Figure 2. Looking dynamics across social-cognition tasks. Panel A shows a task-level descriptive comparison between the full-face facial emotion-recognition task (TREC) and the eye-region mental-state attribution task (RMET), including mean fixation count, cumulative fixation duration, and reaction time. Panel B shows TREC gaze allocation by facial area of interest, with fixation count and cumulative fixation duration for the eyes, nose, and mouth in the total sample; data are presented as mean ± SD, and brackets indicate post hoc contrasts reported for the total sample. Panel C presents an effect-size heatmap for TREC facial-region contrasts, showing Cohen’s *d* values for fixation count and cumulative fixation duration; positive values indicate higher values for the first facial region named in each contrast. Panel D shows TREC gaze laterality by performance group, with fixation count on the left and right hemiface in participants scoring above the sample mean versus those scoring at or below the sample mean; data are presented as mean ± SD, and asterisks indicate significant between-group differences at each hemiface. ***p* < .01, ****p* < .001. TREC, *Test de Reconocimiento Emocional en Caras*; RMET, *Reading the Mind in the Eyes Test*; AOI, area of interest.

Table 1. Sample characteristics by sex.

	Total (N = 19)		Men (n = 11)		Women (n = 8)		t	p
	Mean	SD	Mean	SD	Mean	SD		
Age (years)	35.89	11.35	37.10	13.95	34.25	6.92	-0.58	.57
Education (years)	13.16	2.32	13.00	2.57	13.38	2.07	0.35	.73
TREC score	18.00	1.49	17.82	1.66	18.25	1.28	0.61	.55
RMET score	25.10	2.20	24.55	1.92	25.88	2.47	1.32	.20
Right-handedness (%)	89.47		90.91		87.50			

Table 2. TREC fixation count and cumulative fixation duration by facial area of interest.

Outcome	Sample	Eyes (E)		Nose (N)		Mouth (M)		Cohen’s <i>d</i>			post hoc
		Mean	SD	Mean	SD	Mean	SD	Eyes vs Nose	Eyes vs Mouth	Nose vs Mouth	
Fixation count	Total sample	89.35	51.40	83.26	32.96	21.43	16.84	0.14	1.78	2.36	E > M: <i>t</i> = -17.58, <i>p</i> < .001; N > M: <i>t</i> = -16.00, <i>p</i> < .001

Outcome	Sample	Eyes (E)		Nose (N)		Mouth (M)		Cohen's <i>d</i>			post hoc
		Mean	SD	Mean	SD	Mean	SD	Eyes vs Nose	Eyes vs Mouth	Nose vs Mouth	
Fixation duration (ms)	Men	101.33	51.89	87.74	35.87	21.33	20.81	0.30	2.02	2.26	
	Women	72.88	49.11	77.11	29.68	21.57	10.51	-0.10	1.44	2.49	
	Total sample	210.62	205.21	173.11	115.84	49.60	50.35	0.22	1.08	1.38	E > M: $t = -13.94, p < .001$; N > M: $t = -10.70, p < .001$
	Men	221.41	122.81	186.47	90.80	54.87	58.56	0.32	1.73	1.72	
	Women	195.72	93.81	154.81	48.54	42.35	38.91	0.18	0.73	1.04	

Table 3. TREC gaze laterality by performance group.

	TREC > 18		TREC ≤ 18		Between-group	
	Mean	SD	Mean	SD	<i>t</i>	<i>p</i>
Left hemiface fixations	154.47	64.63	33.68	31.37	-5.42	< .001
Right hemiface fixations	70.84	56.14	137.63	39.02	3.07	.007

Table 4. RMET task-level gaze metrics by sex.

	Total sample		Men		Women		<i>p</i>	Cohen's <i>d</i>
	Mean	SD	Mean	SD	Mean	SD		
Number of fixations	467.16	171.99	508.98	165.55	409.65	174.30	> .05	0.58
Fixation duration (ms)	2215.87	1346.71	2432.10	1276.39	1918.55	1470.27	> .05	0.37

4. Discussion

This pilot study examined whether eye-tracking can provide interpretable, task-dependent information about visual sampling during computerized social-cognition assessment. The findings should be interpreted as preliminary and hypothesis-generating, given the small sample size, exploratory analytical approach, and absence of confirmatory correction for multiple testing. From this perspective, the results suggest that gaze-derived metrics may yield coherent patterns of visual sampling, but that their meaning depends critically on task structure, stimulus layout, task demands, and area-of-interest definition[21].

The clearest pattern emerged in the full-face facial emotion-recognition task, where gaze was preferentially allocated to the eyes and nose, with markedly less sampling of the mouth. This pattern should not be interpreted as simple dominance of the eye region, but rather as structured allocation of attention across facial regions with different potential diagnostic value. Classical models of face perception propose that facial processing depends on distributed systems encoding both invariant and changeable facial information[4]. Neuropsychological evidence also indicates that the eye region can be especially informative for emotion recognition[26]. More recent work suggests that superior face-recognition performance may depend not only on the amount of information sampled, but also on the computational value of the sampled facial information[27]. Within this framework, the present pattern may reflect a strategy in which participants combine eye-region sampling with more central

facial integration, although this remains a hypothesis because facial diagnosticity, emotional category, and stimulus salience were not experimentally manipulated.

The association between TREC performance and eye-region sampling is consistent with the idea that more accurate performance may be supported by attention to diagnostically informative facial features. Previous eye-tracking work has reported positive associations between expression-recognition performance and attention to the eyes[28]. However, the present findings must be interpreted cautiously. The performance groups were defined using a sample-derived cut-off, the sample was small, and the analyses were exploratory. It therefore remains unclear whether greater eye-region sampling reflects a more efficient perceptual strategy, increased task engagement, slower evidence accumulation, or compensatory processing. Disentangling these possibilities will require larger samples, trial-level modelling, preregistered contrasts, and continuous performance measures, because fixation-derived metrics require task-specific interpretation rather than generic interpretation as direct markers of attention or cognition[14,21].

The laterality findings raise an additional hypothesis regarding the spatial organization of visual sampling. Participants with higher performance showed greater fixation on the left hemiface, whereas those with lower performance showed greater fixation on the right hemiface. This pattern is compatible with previous evidence suggesting lateralized biases in face processing[29]. It is also broadly compatible with models proposing right-hemisphere involvement in processing changeable and socioemotional facial cues[4]. However, the present data do not allow discrimination between neurocognitive, stimulus-related, or methodological explanations. Laterality effects may depend on stimulus composition, emotional category, segmentation procedures, or display characteristics, and should therefore be considered exploratory signals rather than evidence of a stable mechanism.

The RMET yielded a different profile of results. Compared with the TREC, it elicited higher fixation count, longer cumulative fixation duration, and longer response times. This pattern may reflect more prolonged inspection under conditions of restricted perceptual input, consistent with the structure of the RMET as an eye-region task requiring mental-state attribution from limited visual information[24]. However, because the RMET stimulus is restricted to the eye region, it does not allow spatially resolved analysis of gaze allocation across facial features. The higher number and duration of fixations should therefore not be interpreted straightforwardly as better, worse, or more socially informative scanning. This distinction is important because the RMET is widely used as a measure of theory of mind, but recent psychometric work has questioned whether it indexes a unitary construct, suggesting instead that performance may reflect heterogeneous perceptual, lexical, semantic, and inferential components[16,17]. The present findings reinforce the broader methodological point that the same eye-tracking metric can have different meanings across tasks, depending on the perceptual information made available by the paradigm[14,21].

The absence of clear sex differences in gaze allocation should also be interpreted cautiously. The present study was not powered to detect sex effects, and the male and female subsamples were small. Previous work has reported sex-related differences in facial-expression recognition and face scanning, including greater attention to the eyes in women in some samples[28]. Therefore, the present null pattern should not be interpreted as evidence that sex plays no role in social-cognitive visual sampling. Rather, it indicates that this pilot sample did not provide reliable support for sex-related differences in gaze allocation. Similarly, the observed associations with age and emotional valence should be considered exploratory. Age-related differences in face scanning and emotion recognition have been documented previously[30]. Emotional salience and context can also shape how faces are inspected and interpreted[30]. However, these effects were not the primary focus of the present study and were not corrected for multiple comparisons.

From a methodological perspective, the main contribution of the study is to illustrate that eye-tracking can be integrated into computerized social-cognition tasks and can yield interpretable patterns of visual sampling, provided that those patterns are understood in relation to task design. Eye-tracking has been described as a fine-grained process-tracing method for cognitive and affective mechanisms[21]. However, area-of-interest analyses require careful definition, reporting, and

interpretation because AOI-based results depend on spatial accuracy, stimulus layout, and how gaze samples are assigned to regions of interest[14]. In this study, the TREC provided spatially resolved information about gaze allocation across facial features, whereas the RMET provided global indices of inspection under perceptual constraint. These differences do not indicate that one task is superior to the other. Rather, they suggest that full-face emotion-recognition tasks are better suited to studying spatial allocation across facial features, whereas eye-region tasks are better suited to examining inspection time, response latency, and inference under restricted perceptual input[17,24].

The study also has implications for future clinical research, although no clinical claims can be drawn from the present data. Social-cognitive impairments are well established across schizophrenia-spectrum disorders[31]. Recent clinical frameworks continue to emphasize the need for more precise and process-oriented assessment of social cognition[31]. Eye-tracking may contribute to this effort by providing information about how social stimuli are sampled, rather than only whether they are correctly interpreted[21]. However, the present study does not demonstrate diagnostic validity, predictive value, or biomarker potential. Instead, it provides a methodological basis for future studies examining whether altered social cognition reflects differences in perceptual sampling, attentional prioritization, affective salience, or compensatory visual strategies.

Strengths & limitations

A key strength of this pilot study is its process-level design: eye-tracking was embedded in computerized social-cognition tasks to complement accuracy scores with information on how visual information was sampled during task performance. The use of two tasks with different perceptual constraints is also informative. The full-face TREC enabled spatially resolved analysis of gaze allocation across facial regions and hemifields, whereas the eye-region RMET provided global inspection metrics under restricted visual input. This contrast clarifies that gaze-derived measures are not task-independent markers, but metrics whose meaning depends on stimulus structure and inferential demand.

The study also has limitations that define its inferential scope. The sample was small, as expected in a pilot study, so findings should be interpreted as preliminary signals rather than normative patterns. Analyses were exploratory, uncorrected for multiple comparisons, and partly based on sample-derived cut-offs; future studies should use larger samples, preregistered hypotheses, continuous models, and trial-level analyses. Calibration difficulties may have affected spatial precision, especially for area-of-interest measures. The healthy adult sample precludes clinical or diagnostic conclusions, and static stimuli limit ecological generalization. These limitations do not diminish the study's value, but specify its role: a feasibility and signal-characterization step to identify task-dependent gaze metrics for later validation.

Further directions

Future studies should validate these preliminary gaze patterns in larger, preregistered samples using continuous, trial-level models that jointly examine fixation count, cumulative fixation duration, reaction time, accuracy, emotional valence, and task format. This would help determine whether the observed patterns reflect stable visual-sampling strategies or sample-specific exploratory signals. Methodologically, future work should refine the interpretation of gaze metrics by improving calibration, standardizing area-of-interest procedures, and testing the reliability of each eye-tracking index. Direct comparisons between full-face, eye-region, dynamic, and context-rich stimuli are needed to clarify which measures are task-specific and which generalize across social-cognitive paradigms. Clinically, this approach should be extended to populations with heterogeneous social-cognitive difficulties, including schizophrenia-spectrum disorders, autism, mood disorders, and neurocognitive disorders. Eye-tracking should be tested as a process-level complement to behavioral accuracy, with the aim of identifying whether altered performance reflects atypical facial sampling, lateralized exploration, emotional-valence sensitivity, slowed inference, or compensatory visual strategies.

5. Conclusion

This pilot study suggests that eye-tracking can be feasibly integrated into computerized social-cognition assessment and may provide task-dependent information about visual sampling. Full-face emotion recognition allowed spatially resolved characterization of gaze allocation across facial regions, whereas eye-region mental-state inference mainly yielded global inspection metrics. These findings are preliminary and do not establish stable gaze profiles, mechanisms, diagnostic utility, or clinical markers. Rather, they provide a methodological basis for larger studies testing whether task-sensitive gaze-derived measures can contribute to the characterization of social-cognitive processing.

Author Contributions: Conceptualization, R.A.-A., S.O.-R. and A.D.-P.; methodology, M.S.-R. and L.R.-C.; software, L.R.-C.; investigation, R.A.-A., M.S.-R., E.S.-S. and A.D.-P.; data curation, M.S.-R. and L.R.-C.; formal analysis, M.S.-R. and L.R.-C.; visualization, L.R.-C. and A.D.-P.; interpretation of results, R.A.-A. and A.D.-P.; writing—original draft preparation, R.A.-A. and A.D.-P.; writing—review and editing, R.A.-A., M.S.-R., E.S.-S., L.R.-C., S.O.-R. and A.D.-P.; supervision, R.A.-A. and A.D.-P.; project administration, R.A.-A.; funding acquisition, S.O.-R. and R.A.-A. All authors have read and agreed to the submitted version of the manuscript.

Funding: This project was funded by Ambar Telecomunicaciones SL under the CogniBio project, supported by the Innova 2019 Programme of the Regional Ministry of Innovation, Industry, Tourism and Trade of the Government of Cantabria, and by a Miguel Servet contract awarded to Dr. Rosa Ayesa-Arriola by the Instituto de Salud Carlos III (CP18/00003). No pharmaceutical company provided financial support for this study.

Acknowledgments: The authors would like to thank the Photonic Engineering Group at the University of Cantabria for their technological support, Ambar Telecomunicaciones SL for their contribution to the CogniBio project, and all volunteers who participated in this pilot study. The authors also acknowledge the IDIVAL Mental Illness Research Department for its support in participant recruitment, study coordination, and research infrastructure.

References

1. Adolphs R (2009) The social brain: neural basis of social knowledge. *Annu Rev Psychol* 60:693–716. <https://doi.org/10.1146/annurev.psych.60.110707.163514>
2. Arioli M, Crespi C, Canessa N (2018) Social Cognition through the Lens of Cognitive and Clinical Neuroscience. *Biomed Res Int* 2018:4283427. <https://doi.org/10.1155/2018/4283427>
3. Bölte S (2025) Social cognition in autism and ADHD. *Neuroscience & Biobehavioral Reviews* 169:106022. <https://doi.org/10.1016/j.neubiorev.2025.106022>
4. Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4:223–233. [https://doi.org/10.1016/s1364-6613\(00\)01482-0](https://doi.org/10.1016/s1364-6613(00)01482-0)
5. Pitcher D (2025) Neuropsychological evidence of a third visual pathway specialized for social perception. *Nat Commun* 16:5774. <https://doi.org/10.1038/s41467-025-61396-8>
6. Rosati AG (2026) Cognitive adaptations in humans and other great apes. 384–396. <https://doi.org/10.1016/B978-0-443-27380-3.00054-3>
7. Shafiei M, Reik M, Görner M, et al (2026) Rhesus monkeys use both eye and head gaze to reallocate covert spatial attention facilitating visual perception. *Cogn Affect Behav Neurosci*. <https://doi.org/10.3758/s13415-025-01383-0>
8. Frith CD, Frith U (2007) Social cognition in humans. *Curr Biol* 17:R724–32. <https://doi.org/10.1016/j.cub.2007.05.068>
9. Cavallo ND, Giacobbe C, Baiano C, et al (2026) Neural correlates of social cognition in stroke and traumatic brain injury: A systematic review. *Cogn Affect Behav Neurosci*. <https://doi.org/10.3758/s13415-026-01438-w>
10. Bohn M, Prein JC, Ayikoru A, et al (2026) A universal of human social cognition: Children from 17 communities process gaze in similar ways. *Child Dev* 97:219–232. <https://doi.org/10.1093/chidev/aacaf017>
11. Caldara R (2017) Culture Reveals a Flexible System for Face Processing. *Curr Dir Psychol Sci* 26:249–255. <https://doi.org/10.1177/0963721417710036>

12. Hessels RS, Iwabuchi T, Niehorster DC, et al (2025) Gaze behavior in face-to-face interaction: A cross-cultural investigation between Japan and The Netherlands. *Cognition* 263:106174. <https://doi.org/10.1016/j.cognition.2025.106174>
13. Alkan N (2025) Recognition and Misclassification Patterns of Basic Emotional Facial Expressions: An Eye-Tracking Study in Young Healthy Adults. *J Eye Mov Res* 18:53. <https://doi.org/10.3390/jemr18050053>
14. Hooge ITC, Nyström M, Niehorster DC, et al (2026) The fundamentals of eye tracking part 6: Working with areas of interest. *Behav Res* 58:65. <https://doi.org/10.3758/s13428-025-02937-3>
15. De Biase R, Ponari M, Sagliano L (2026) I have my eye on the emotion: a review of eye-tracking studies exploring processing of emotional facial expressions. *Psychological Research* 90:80. <https://doi.org/10.1007/s00426-026-02294-6>
16. Higgins WC, Kaplan DM, Deschrijver E, Ross RM (2025) Why most research based on the Reading the Mind in the Eyes Test is unsubstantiated and uninterpretable: a response to Murphy and Hall (2024). *Clinical Psychology Review* 115:1–5. <https://doi.org/10.1016/j.cpr.2024.102530>
17. Hafner RM, Johnson BN, Tone EB, et al (2026) A multi-site psychometric evaluation of the Reading the Mind in the Eyes Test – Revised. *Personality and Individual Differences* 250:113526. <https://doi.org/10.1016/j.paid.2025.113526>
18. Corbera S, Kurtz MM, Achim AM, et al (2025) International perspective on social cognition in schizophrenia: current stage and the next steps. *European Psychiatry* 68:e9. <https://doi.org/10.1192/j.eurpsy.2024.1776>
19. Dodich A, Panzavolta A, Funghi G, et al (2025) International consensus for the assessment of social cognition in neurocognitive disorders: framework definition and clinical recommendations of the SIGNATURE initiative. *Alzheimers Res Ther* 18:6. <https://doi.org/10.1186/s13195-025-01908-2>
20. Yang Q, Fu Y, Yang Q, et al (2024) Eye movement characteristics of emotional face recognizing task in patients with mild to moderate depression. *Front Neurosci* 18:. <https://doi.org/10.3389/fnins.2024.1482849>
21. Rahal R-M, Fiedler S (2019) Understanding cognitive and affective mechanisms in social psychology through eye-tracking. *Journal of Experimental Social Psychology* 85:103842. <https://doi.org/10.1016/j.jesp.2019.103842>
22. Baron-Cohen S, Wheelwright S, Jolliffe T (1997) Is there a “language of the eyes”? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition* 4:311–331. <https://doi.org/10.1080/713756761>
23. Huerta-Ramos E, Ferrer-Quintero M, Gómez-Benito J, et al (2021) Translation and validation of the Baron-Cohen Face Test in the Spanish population. *Actas Esp Psiquiatr* 49:106–113
24. Baron-Cohen S, Wheelwright S, Hill J, et al (2001) The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry* 42:241–251
25. Fernández-Abascal EG, Cabello R, Fernández-Berrocal P, Baron-Cohen S (2013) Test-retest reliability of the ‘Reading the Mind in the Eyes’ test: a one-year follow-up study. *Molecular Autism* 4:33. <https://doi.org/10.1186/2040-2392-4-33>
26. Adolphs R, Gosselin F, Buchanan TW, et al (2005) A mechanism for impaired fear recognition after amygdala damage. *Nature* 433:68–72. <https://doi.org/10.1038/nature03086>
27. Dunn JD, Varela V, Popovic B, et al (2025) Super-recognizers sample visual information of superior computational value for facial recognition. *Proc Biol Sci* 292:20252005. <https://doi.org/10.1098/rspb.2025.2005>
28. Hall JK, Hutton SB, Morgan MJ (2010) Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? *Cognition and Emotion* 24:629–637. <https://doi.org/10.1080/02699930902906882>
29. Parente R, Tommasi L (2008) A bias for the female face in the right hemisphere. *Laterality* 13:374–386. <https://doi.org/10.1080/13576500802103495>
30. Noh SR, Isaacowitz DM (2013) Emotional Faces in Context: Age Differences in Recognition Accuracy and Scanning Patterns. *Emotion* 13:238–249. <https://doi.org/10.1037/a0030234>

31. Green MF, Horan WP, Lee J (2015) Social cognition in schizophrenia. *Nat Rev Neurosci* 16:620–631.
<https://doi.org/10.1038/nrn4005>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.