Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.37735

Social Cognitive Processing Deficits in Adolescents with Social Anxiety: Dynamic Modeling of Gaze Avoidance in Virtual Peer Contexts

Duanyue Qiu

Institute of Psychology, Shanghai Normal University, Shanghai, China
duanyueqiu02550@outlook.com

Abstract. Adolescent social anxiety is frequently associated with atypical visual attention during peer evaluation, while static stimuli and aggregated gaze measures provide limited information about temporal transitions among threat vigilance, disengagement difficulty, and sustained avoidance. Virtual peer interactions involving acceptance, neutral, and rejection feedback were therefore developed. Gaze position, pupil variation, and head pose were collected from 60 adolescents using the HTC Vive Pro Eye, while Hidden Markov Models and an Elastic Net classifier were used to characterize latent attentional states. Under rejection feedback, eye-region fixation was 31.6% in the social anxiety group and 42.8% in healthy controls. Return latency to the eye region was 2.08 s and 1.43 s, respectively, with a significant group-by-feedback interaction. HMM analysis showed a between-group transition difference of 0.16 from eye locking to environmental avoidance and −0.13 from environmental avoidance back to social engagement. The dynamic-feature classifier achieved an AUC of 0.84, compared with 0.71 for the static gaze model. These findings indicate that gaze avoidance in adolescent social anxiety is temporally structured and context dependent. Continuous state modeling provides a more complete representation of social-cognitive processing abnormalities during peer evaluation and offers a computational basis for objective behavioral assessment.

Keywords: Social Anxiety, Adolescents, Virtual Peers, Eye Tracking, Hidden Markov Model

1. Introduction

This stage of life is marked by increased susceptibility to evaluations by peers, thus the relevance of the concept of social anxiety in the context of social participation, friendship, and expression of oneself. Eye-tracker studies revealed gaze avoidance among socially anxious adolescents during dynamic social interactions [1]. Such a trend implies that visual attention could serve as an observable marker of social-cognitive functioning. Socially anxious adolescents would also take more time to disengage from other people's gazes, thus implying that difficulties in visual attention precede avoidance behaviors [2]. Social-communicative features influence the process of distribution of visual attention during the interpersonal communication [3]. All of these findings allow for seeing gaze behavior as a dynamic sequence, rather than one specific attentional

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.37735

impairment. In addition, static faces and aggregate measures of fixation might hide the dynamics between vigilance, difficulty in disengagement, and avoidance. This methodological approach involves adolescents taking part in virtual peer interactions with acceptance, neutral, and rejecting feedback, while gaze position, pupil fluctuation, and head orientation will be measured. Latent gaze states and transitions will be estimated using Hidden Markov Models, and thus allow for analysis of avoidance in different types of feedback, interaction stages, and anxiety level.

2. Literature review

2.1. Social-evaluative processing

Evaluation-based social processing in the experience of social anxiety includes anticipated evaluation, self-focus, and protection during social interactions. Constrained visual scanpaths in emotion recognition imply that socially anxious teens process visual information from the face less adaptively, thus possibly limiting their ability to interpret social cues [4]. Increased monitoring is observable early in the social situation, slowly changing as a result of increased evaluation. The research on dynamic social threat has demonstrated that attention can be directed away from socially relevant areas as a consequence of increased evaluative stress [5]. This order illustrates the possibility of threat monitoring and avoidance within the same interaction. The fear of being negatively evaluated can increase sensitivity to ambiguous peer feedback, and the state of anxiety can determine when disengagement starts.

2.2. Gaze-avoidance dynamics

Gaze avoidance should not be conceptualized solely in terms of a decreased level of eye-region fixations, but rather as a temporal phenomenon. Using hidden Markov modeling, it has been established that socially anxious people may find themselves in different attentional states, which an average obscures [6]. The first-fixation latency, duration of the fixation, size of the saccade and changes from facial to environmental regions may be considered as interconnected components of a single sequence. Gaze avoidance patterns have been found to differ among socio-evaluative tasks in virtual reality studies employing eye tracking [7], thus proving that gaze avoidance is interactionally dependent. The vigilance-to-avoidance trajectory would imply that detection of the threat and disengagement as a result of safety considerations occur at different points in time.

2.3. Virtual interaction modeling

Virtual peer settings enable a controllable social evaluation experience while keeping the face expression, body postures, spatial gaze, and behavior responses intact. The studies done among adolescents have revealed that anxiety among peer presence is associated with fear of negative evaluations [8], which is important for the significance of social threat in specific context. Virtual settings allow for standardization of acceptance, neutrality, and rejection while keeping the interpersonal cues intact. Mobile eye tracking of adolescent social performance in realistic conditions has shown that the directed attention of gaze can be used as a predictor for anxiety symptoms [9]. This helps in increasing ecological validity but the summary statistics reduce the continuous behavior into a single average value.

DOI: 10.54254/2755-2721/2026.37735

3. Experimental methods

3.1. Participant recruitment

An a priori power analysis was conducted using G Power. The primary test focused on the repeated measures interaction between two groups across three peer feedback conditions. A medium effect size, $\alpha = 0.05$, and statistical power of 0.80 were specified. Based on this calculation, 60 valid participants were required. Considering potential eye tracking calibration failure, VR discomfort, and participant withdrawal, approximately 70 adolescents aged 13 to 17 were initially recruited. The social anxiety group was recruited through child and adolescent mental health clinics. After confirmation using the MINI KID, 30 participants were retained. The control group was recruited through cooperating schools, with 30 adolescents matched by age and sex. All eye tracking data were collected by the researchers in a laboratory equipped with the HTC Vive Pro Eye. No public dataset was used. Screening covered social anxiety, fear of negative evaluation, depressive symptoms, visual function, and psychiatric medication use. Participants with recent medication dosage changes, severe depression, or other major psychiatric disorders were excluded. For participants receiving stable medication, medication type, dosage, and treatment duration were recorded. Mild depressive symptoms and stable medication status were included as covariates in subsequent models to reduce confounding effects from comorbidity and treatment.

3.2. Virtual peer task

Before the formal experiment, an additional 10 to 12 adolescents completed a pilot study and were not included in the final sample. The pilot test served to optimize the eye-tracking calibration, length of trial, feedback strength, virtual reality discomfort, and HMM time window. The experimental task was performed in the virtual classroom environment developed using Unity 2022.3. In total, 36 trials were conducted by each participant: 12 trials for accepting, 12 neutral, and 12 rejecting. They were randomly ordered into six balanced blocks. Manipulation of the feedback was carried out through the virtual peer's face expression, voice tone, and posture. Direct gaze stayed the same in order not to contaminate the measure of gaze avoidance with gaze shifts due to stimuli [10]. Binocular gaze coordinates, pupil size, and head pose were measured at the rate of 120 Hz by the HTC Vive Pro Eye. Areas of interest were eyes, other parts of the face, the torso, and the background. Each trial consisted of baseline stage, direct gaze stage, evaluative feedback stage, and verbal response stage. At the end of each block, the participants had to rate their state anxiety from 0 to 10 through the VR interface. Automatic recording of the verbal response latency from the end of feedback to the beginning of the first valid vocal response by the headset microphone was provided.

3.3. Gaze-state modeling

Continuous eye tracking data were processed in Python 3.11 through temporal synchronization, blink removal, interpolation of short missing segments, and velocity threshold based fixation detection. The time window determined in the pilot study was then applied to extract AOI location, fixation duration, saccade amplitude, pupil variation, and head deviation [11]. The Hidden Markov Model was first used to describe the probabilistic relationship between observable gaze sequences and latent attentional states, as shown in Equation 1.

DOI: 10.54254/2755-2721/2026.37735

$$P(O,S|\lambda)=\pi_{s_1} b_{s_1}(o_1)\prod_{t=2}^{T} a_{s_{t-1}s_t} b_{s_t}(o_t) \tag{1}$$

Here, O denotes the continuous gaze observation sequence, S denotes the latent state sequence, π represents the initial state probability, a_{ij} represents the transition probability from state i to state j, and b represents the observation emission probability.

Models containing three to six latent states were then compared using BIC, as shown in Equation 2.

$$BIC=-2\ln(\widehat{L})+k\ln(N) \tag{2}$$

Here, $\widehat{L}$ denotes the maximum likelihood, k denotes the number of free parameters, and N denotes the number of valid observations. The model with the lowest BIC was used to extract state occupancy, mean dwell time, and transition probabilities. These dynamic features were then entered into an Elastic Net logistic classifier. Its objective function is shown in Equation 3.

$$\min_{\beta_0,\beta}\left[-\frac{1}{n}\sum_{i=1}^{n}(y_i\ln p_i+(1-y_i)\ln(1-p_i))+\lambda\left(\alpha\|\beta\|_1+\frac{1-\alpha}{2}\|\beta\|_2^2\right)\right] \tag{3}$$

Here, y_i denotes the group label, p_i denotes the predicted probability of belonging to the social anxiety group, β denotes the feature coefficient, λ denotes the regularization strength, and α controls the relative contribution of L1 and L2 penalties. All classification splits were performed at the participant level using nested five fold cross validation. This prevented different trials from the same participant from appearing in both the training and test sets.

4. Results

4.1. Behavioral gaze patterns

After quality control, 2,048 of 2,160 trials were retained, corresponding to a valid-trial rate of 94.8%. Under rejection feedback, adolescents with social anxiety showed a lower eye-region fixation proportion than healthy controls, 31.6% ± 8.4% versus 42.8% ± 7.9%, together with greater background fixation, 29.4% ± 7.6% versus 18.7% ± 6.9%. Their return latency to the eye region was also longer, 2.08 ± 0.61 s compared with 1.43 ± 0.48 s. State anxiety increased to 6.7 ± 1.5 in the social anxiety group, while the control group reached 3.1 ± 1.2. The group-by-feedback interaction for eye-region fixation was significant, $F(2,116) = 6.84$, $p = 0.002$, indicating that rejection feedback produced a stronger shift away from socially salient facial information in socially anxious adolescents.

DOI: 10.54254/2755-2721/2026.37735

Table 1. Behavioral gaze measures across peer feedback conditions

Measure	Acceptance SA	Acceptance HC	Neutral SA	Neutral HC	Rejection SA	Rejection HC
Eye-region fixation proportion (%)	38.9 ± 8.1	45.1 ± 7.4	35.7 ± 7.9	43.8 ± 7.6	31.6 ± 8.4	42.8 ± 7.9
Return latency to eye region (s)	1.71 ± 0.52	1.39 ± 0.44	1.84 ± 0.57	1.41 ± 0.46	2.08 ± 0.61	1.43 ± 0.48
Background fixation proportion (%)	22.4 ± 6.8	18.1 ± 6.4	25.1 ± 7.1	18.4 ± 6.6	29.4 ± 7.6	18.7 ± 6.9
State anxiety score	5.3 ± 1.4	2.8 ± 1.1	5.8 ± 1.5	2.9 ± 1.1	6.7 ± 1.5	3.1 ± 1.2
Verbal response latency (s)	1.62 ± 0.48	1.25 ± 0.37	1.78 ± 0.51	1.31 ± 0.39	2.05 ± 0.56	1.36 ± 0.41

4.2. Latent-state classification

Among the three- to six-state HMM candidates, the four-state model produced the lowest BIC and was retained for further analysis. Under rejection feedback, environmental-avoidance occupancy reached 0.34 in the social anxiety group compared with 0.21 in controls. As shown in Figure 1, the transition probability from eye locking to environmental avoidance was 0.37 versus 0.21, yielding $\Delta p = 0.16$, 95% CI [0.08, 0.24], $q = 0.006$. In contrast, the transition from environmental avoidance back to social engagement was lower in the social anxiety group, 0.18 versus 0.31, with $\Delta p = -0.13$, 95% CI [-0.21, -0.05], $q = 0.011$. The Elastic Net classifier based on dynamic HMM features achieved an AUC of 0.84, macro-F1 of 0.79, sensitivity of 0.80, and specificity of 0.77. The static-gaze baseline model achieved an AUC of 0.71, indicating that transition-based features provided stronger discrimination between groups.

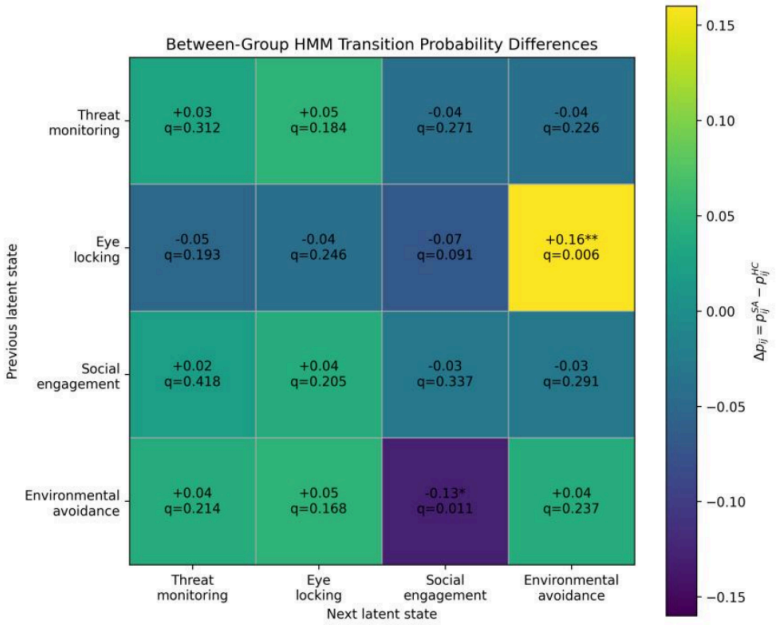


Figure 1. Between group HMM latent state transition difference matrix

DOI: 10.54254/2755-2721/2026.37735

5. Discussion

Reduced eye-region fixation and prolonged return latency after rejection suggest that atypical attention in socially anxious adolescents reflects disengagement from social information following evaluative threat rather than a constant reduction in social attention. The HMM results clarify this temporal structure. Increased transitions from eye locking to environmental avoidance, together with fewer transitions from avoidance back to social engagement, indicate a sequence from early threat monitoring to later safety-oriented avoidance. This pattern differs from static comparisons based on total fixation duration and may help explain previous inconsistencies between vigilance and avoidance findings. The increase in AUC from 0.71 for static gaze measures to 0.84 for dynamic features further indicates that transition patterns contain temporal information lost through aggregation. Elevated state anxiety under rejection feedback also links gaze reallocation with perceived evaluative pressure. Social gaze behavior should therefore be interpreted in relation to interaction phase and feedback valence.

6. Conclusion

Virtual peer interaction, head-mounted eye tracking, and dynamic probabilistic modeling provide an integrated framework for examining social-cognitive processing in adolescent social anxiety. Comparisons among 60 adolescents showed reduced eye-region fixation, prolonged attentional return, and stronger transitions toward environmental avoidance after rejection feedback in the social anxiety group. HMM analysis further differentiated threat monitoring, eye locking, social engagement, and environmental avoidance, while Elastic Net classification demonstrated the discriminative value of these temporal features. The overall findings show that gaze avoidance cannot be adequately represented by average fixation duration alone. Important differences emerge from how attentional states develop, persist, and transition over time. Combining controlled virtual interaction with computational gaze analysis therefore provides a feasible approach for understanding social-cognitive deficits in adolescent social anxiety and developing objective behavioral markers for computer-assisted assessment.

References

[1] Chen, J., et al. (2023). Social anxiety is related to reduced face gaze during a naturalistic social interaction. *Anxiety, Stress, & Coping, 36*(4), 460–474.
[2] Niu, Y., et al. (2023). Context and domain matter: The error-related negativity in peer presence predicts fear of negative evaluation, not global social anxiety, in adolescents. *Psychological Medicine, 53*(14), 6899–6909.
[3] Zeka, F., et al. (2023). Examination of gaze behaviour in social anxiety disorder using a virtual reality eye-tracking paradigm: Protocol for a case–control study. *BMJ Open, 13*(8), e071927.
[4] Lidle, L. R., & Schmitz, J. (2024). Assessing visual avoidance of faces during real-life social stress in children with social anxiety disorder: A mobile eye-tracking study. *Child Psychiatry & Human Development, 55*(1), 24–35.
[5] Rubin, M., et al. (2024). Attentional heterogeneity in social anxiety disorder: Evidence from hidden Markov models. *Behaviour Research and Therapy, 173*, 104461.
[6] Chen, J., van den Bos, E., & Westenberg, P. M. (2024). Does gaze anxiety predict actual gaze avoidance and is it more informative than social anxiety? *Journal of Behavior Therapy and Experimental Psychiatry, 82*, 101896.
[7] Howell, A. N., et al. (2024). Mutual eye gaze and vocal pitch in relation to social anxiety and depression: A virtual interaction task. *Journal of Affective Disorders, 363*, 282–291.
[8] Azriel, O., et al. (2024). Attention bias vs. attention control modification for social anxiety disorder: A randomized controlled trial. *Journal of Anxiety Disorders, 101*, 102800.
[9] Byrne, M. E., et al. (2024). Eye-tracking measurement of attention bias to social threat among youth: A replication and extension study. *Journal of Mood & Anxiety Disorders, 8*, 100075.

DOI: 10.54254/2755-2721/2026.37735

[10] Rauschenbach, A.-L., et al. (2024). Reduced early neural processing of faces in children and adolescents with social anxiety disorder. *Biological Psychology, 191*, 108827.
[11] Hauffe, V., et al. (2025). Early hypervigilance and sustained attention for the eye region in adolescents with social anxiety disorder. *Journal of Anxiety Disorders, 112*, 103016.