Appendix: Understanding Unique Behavioral Patterns through Multimodal Analysis of Eye-Hand Coordination in Autistic Children

Organization of the appendix. In Appendix A, we present a detailed description on how the eye-hand behavioral sequences are collected, including the VR setup and how fixations are extracted. We also provide precise coordinates for the boundaries of different regions on the screen. We provide additional information about the detailed structure of the eye-to-hand translation model, along with the parameters used across the experiments. In Appendix C, we present additional experimental results, along with visualizations.

A Detailed Description of the Dataset

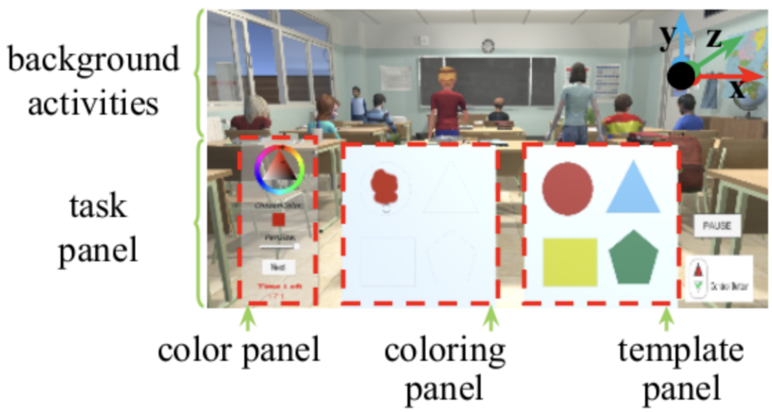


Figure 3: Multimodal virtual classroom interface system

The data was collected through the Multimodal Virtual Classroom Interface system (MVCI) (Yu et al. 2023), shown in Figure 3. The MVCI stimulates an interactive classroom environment, with a digital screen that contains a simple drawing task for the participants. The system incorporates auditory, visual, and tactile stimuli, with moving avatars in the background and bells ringing. The task was completed using a stylus that provided haptic feedback, with friction, weight, drag, and resistance to mimic a real pen. To extract eye-gaze fixations, the dispersion-threshold algorithm (I-DT) was applied, with an angular threshold of $0.7°$ and a minimum fixation duration of 275 ms.

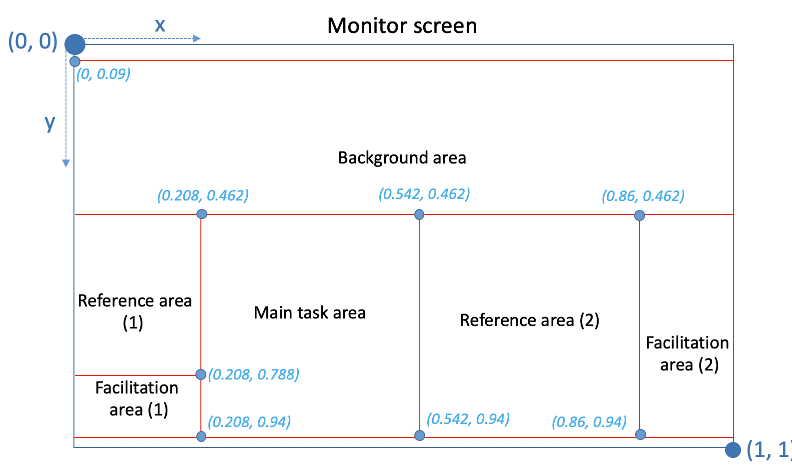
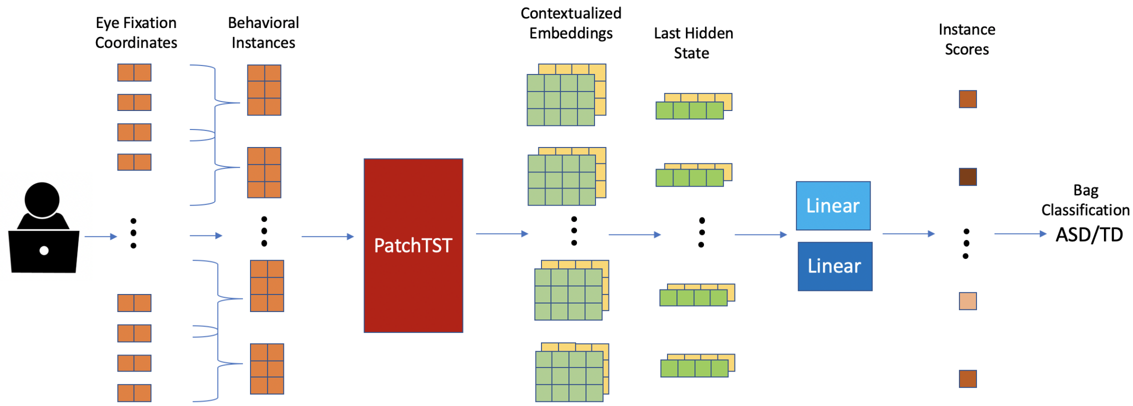


Figure 4: Coordinate boundaries of different regions on the screen which correspond to different tasks or images on the screen.

Participants were between the ages of 11 to 17, and data was collected from 9 autistic and 17 typically developing adolescents. The region boundaries correspond to the different sections on the screen, with coordinates shown in Figure 4

B Additional Details of the Methodology



We prepare our instances from the bag by segmenting each time series sequence into overlapping subsequences of length L timestamps. We use a sliding window with stride d to create these fixation instances $\{x_{i1}, x_{i2}, \ldots, x_{iN}\}$. We then encode each instance with the PatchTST encoder, which we take the final token for a sequence representation as h_{ij}. We then pass our embedding into a MIL layer, which maps it to an instance score s_{ij} for bag B_i. We calculate the bag score as the sum of the top K instances: $\hat{y}_i = \frac{1}{K}\sum s_{ij}$, where s_{ij} are in the top K set. A bag is predicted as ASD if the bag score is positive, meaning positive scores were present in the bag, and TD otherwise. The instances with the highest scores determine the bag label, and thus are the uniquely identifying behaviors. We train the model using the hinge loss, which is defined as $L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i\hat{y}_i)$, where $y_i \in \{-1, 1\}$.

C Additional Experimental Results

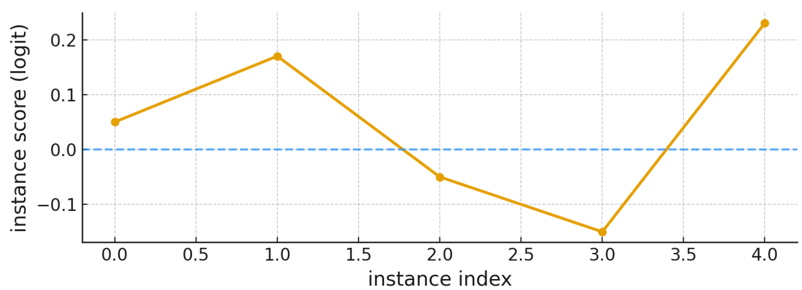


Figure 5: Instance scores corresponding to the fixation sequence of `Individual 1`

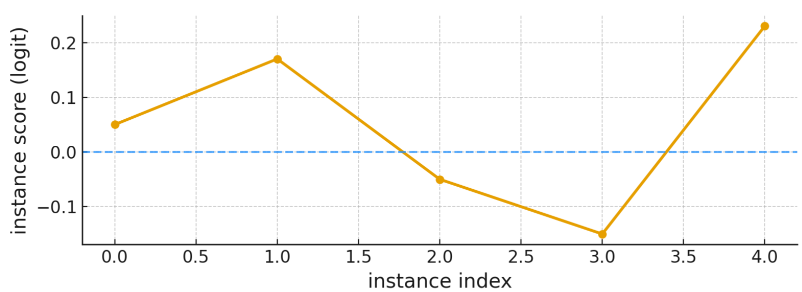


Figure 6: Instance scores corresponding to the fixation sequence of `Individual 1`

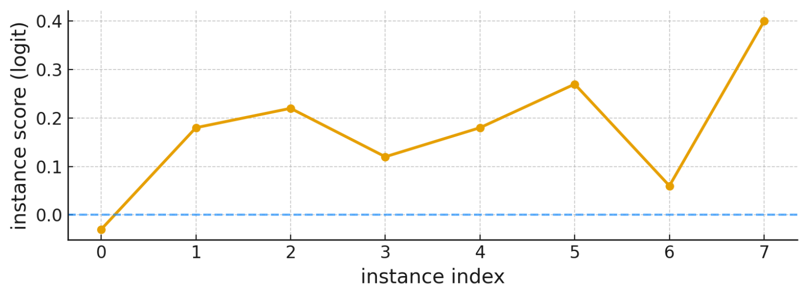


Figure 7: Instance scores corresponding to the fixation sequence of `Individual 2`

This way, we can also pinpoint the time points where unique behavioral patterns occur, which can be further analyzed.

We plot the relationship between instance scores as the time of the sequence progresses. This displays the specific subsequences are most responsible for determining the classification of the sequence, and pinpoints the key moments in the sequence that exhibit strong discriminative behavior. Additionally, we are able to condense segments particularly important for the interpretation on an individual basis. Additionally, we can see behavior fluctuations within a sequence. This provides insight into whether ASD-like behaviors are consistent or appear sporadically throughout the sequence. For example, Figure 6 shows that `Individual 1` exhibits several isolated peaks of signature behavior. In contrast, `Individual 2` as shown in Figure 7 displays a more continuous pattern of atypical behavior. Moreover, this can be used for the classification of high scoring subsequences with known behavioral markers, which can be helpful in various clinical settings.

D Acknowledgments

I would to thank Dr. Zhi Zheng from University of Notre Dame, who served as my mentor during the course of working on this project. Dr. Zheng is an expert in human-computer interaction with years of experience in designing reliable assistive systems for mental health care. Dr. Zheng introduced to me this research area, suggested the overall research topic of using a data-driven approach to understand unique behaviors in children with autism, and shared the dataset that is analyzed in my study. We had weekly meetings to discuss the progress of my work and Dr. Zheng offered important feedback that helped to shape this research. I would also like to thank Zhiwei Yu, who helped preprocess the data by extracting the fixations from the raw data.