EmoBridge: Bridging Emotion Comprehension from Avatars to Human Faces for Individuals with Autism Spectrum Disorder

Jingyun Wang[a], Adam Wynn[b] and Rebecca Wood
Department of Computer Science, Durham University, Stockton Rd, Durham DH1 3LE, UK
jingyun.wang@durham.ac.uk, adam.t.wynn@durham.ac.uk, rebecca.e.wood1@gmail.com

Keywords: Autism Spectrum Disorder, Emotion Recognition, Adaptive Learning, Gamification, Customization.

Abstract: It is challenging for autistic people to interpret and respond appropriately to social scenarios due to differences in emotion understanding and recognition abilities. This paper aims to design, implement, and evaluate an adaptive, gamified tool to support emotion recognition skill development in children with ASD. First, we propose a novel adaptive algorithm that facilitates a personalized progression from avatar to human faces according to the user's real-time performance accuracy. Second, we present the implementation of EmoBridge, a mobile application that integrates this algorithm within a gamified learning session and includes a browse section for users to view emotion-categorized images and upload their own. Third, we report findings from a preliminary experiment involving 18 children aged 7-12, including both neurotypical and ASD-diagnosed participants, and their parents. The results demonstrate EmoBridge's potential to support emotion recognition, particularly for verbal children with ASD, while also revealing specific limitations for non-verbal users. The feedback collected provides valuable insights that will guide future iterations and research, contributing to the development of more inclusive and adaptive assistive technologies.

1 INTRODUCTION

Autism Spectrum Disorder (ASD) is a lifelong developmental disability that affects how individuals communicate and interact with others (National Autistic Society, 2022). It affects people of all ages, genders, and backgrounds, and it is estimated that worldwide, around 1 in 100 children have ASD (Zeidan et al., 2022). Due to different ways of thinking and how ASD individuals experience the world, they find it relatively difficult to empathise with and understand others (NHS Transformation Partners, 2017), and their emotional understanding and recognition abilities differ from neurotypical patterns (Baron-Cohen et al., 2009). Therefore, for autistic people, deciphering the emotional landscape is complex, and often presents significant challenges.

Several applications have been developed to support the emotional understanding of individuals, especially children with ASD. Alves et al. (2013) have implemented a prototype "LIFEisGAME" with five serious game modes using virtual characters. They found that the most difficult task perceived by children with ASD involved identifying the emotion shown on the face of an avatar. They suggested that this is because this task puts the heaviest dependence on the child using their empathy skills, which is a challenge for individuals with ASD. However, their result beyond the interaction with the games was limited to anecdotal responses from parents and therapists. No baseline tests were conducted to compare the skills the children may have developed after using the games.

The Ucime Emocii Intervention Program (Vasilevska Petrovska and Trajkovski, 2019) also developed a cross-platform web application aiming to support autistic individuals in improving their emotional understanding. One game is designed to encourage the understanding of four basic emotions (happy, sad, fear, anger) appearing on human and avatar faces. Thirty-two children with ASD aged 7-15 were split into an experimental group who used the application, and a control group who were not included in any intervention beside the usual educational curriculum. The analysis results suggested that the game had a positive impact on the experimental group. Another key finding is that the highest effect difference was found in the task involving the identification of emotions from avatars, suggesting autistic children find it easier to

[a] https://orcid.org/0000-0001-9325-1789
[b] https://orcid.org/0000-0002-1631-2151

interpret emotions from avatar or pictogram faces than from real human faces. This may be attributed to the reduced perceptual and social complexity of avatar expressions, which present emotions in a more schematic and consistent manner compared to the nuanced variability inherent in real human faces.

Another study which also used both human and avatar faces was carried out to assess the patterns in eye gaze for children with ASD, attempting to recognise the emotions of virtual versus real human faces (Pino et al., 2021). This study involved 29 5-11-year-old children diagnosed with ASD being asked to identify the emotion in images of 8 avatar faces and 8 human faces representing four basic human emotions (happy, surprise, anger and sad). Whilst the participants observed the images, the duration for which their eyes focused on the eyes and mouth of each face was tracked. Results showed that the children looked at the avatar faces much more than the human faces, and the accuracy of emotions identified was also much higher for the avatars. It was also found that different emotions were more effectively depicted by different types of faces. In particular, negative emotions were more easily perceived in the form of the avatars. Therefore, using a combination of both avatars and human faces may be more helpful in developing the emotion recognition skills of children with ASD.

Despite the observed advantages of using avatars, none of the previous studies transferred the skills developed through interaction with avatars into more applicable skills to real-world interactions. Transferable skills are crucial, as in real-world interactions, the children will not be interacting with avatars, and hence avatar-only skills developed will have limited real-world applicability. To address this gap, we pose the following research question: "Can an application which utilises a transition from avatar faces to human faces support individuals with ASD to develop transferable emotion recognition abilities?" Ultimately, this work aims to contribute to a larger vision: empower children with ASD with tools that enhances their understanding of others' emotions, helping them forge stronger relationships and navigate social situations with greater confidence.

The main contribution of this work are: (1) We propose a novel adaptive algorithm that facilitates a transition from avatar faces to real human faces in response to the user's skill level to support individuals with ASD in improving their emotion recognition skills. (2) Building on this algorithm, we developed and deployed a gamified application to deliver a personalized learning path for individuals with ASD. (3) We conducted a study that engaged a diverse participant group, including verbal and non-verbal children with ASD, neurotypical children, and their parents, to collect formative feedback that will inform and shape future research directions.

2 RELATED WORK

2.1 Existing applications for ASD

Autistic people often face challenges in social communication across multiple contexts and may experience a wide array of associated symptoms, including but not limited to sensory sensitivities, cognitive differences, and emotional regulation challenges (NHS Transformation Partners, 2017). These characteristics necessitate personalised support, emphasising the critical role of innovative tools designed specifically for autistic individuals.

Several applications have been developed to support ASD individuals in various aspects of their lives. Novack et al. (2018) evaluated a mobile application designed to cultivate receptive language skills in autistic children. This study underscored the potential of mobile applications as an interactive medium to teach vital skills. Their findings revealed significant improvements in the participants' receptive language capabilities, demonstrating the efficacy of tailored mobile applications. These acquired skills were maintained in the treatment. However, a limitation is that there is no evidence that the acquired skills were transferable to the natural environment.

AutiVid (Aburukba et al., 2017) is an application developed to teach children with ASD some of the basic skills they need in their daily lives by using video modelling. The videos on the application are intended to demonstrate to the child the basic steps of getting ready for school. The application was evaluated by 3 children diagnosed with ASD, where each child was shown a video, and then was asked to complete a task. This process was repeated until they were able to correctly complete the task without assistance. Results found that all the participants were able to complete all tasks after watching the video, and the task that all participants struggled with the most was the one that required the highest number of steps to complete. This study presents positive signs that a video modelling application produces promising results but is limited by user data from a very small sample size and a total of only 4 modules on the application. To conclude that video modelling is a reliable method of intervention, a much wider scope of modules should be tested on a greater and more diverse group of individuals.

Developed to enhance social communication in children with ASD, the MyVoice application (Mohammad and Abu-Amara, 2019) uses a visual framework to teach functional vocabulary and the expression of emotions and needs. Its core functionality enables users to build sentences by combining pictorial symbols, thereby supporting language construction. The users can also add their own images and audio within pre-defined categories. The application was evaluated by one child and two therapists who specialise in working with individuals with ASD. The child quickly adapted, and after just one week of full interaction, was able to express their needs using the application. The therapists also acknowledged the application as efficient and convenient. This shows positive signs that individuals with ASD find mobile applications useful and easy to adapt to, however, the method that was implemented in this study simply digitized a pre-existing physical method of intervention Picture Exchange Communicate System (PECS) (Bondy and Frost, 1994). This may have helped the child to adapt to the application, as they were already familiar with using physical PECS cards to construct sentences and communicate. This is also consistent with the findings of Yee (2012) that existing mobile applications, which digitize PECS interventions, provided an effective way to allow a much more accessible augmentative and alternative communication method for individuals with ASD.

Papoutsi et al. (2018) also explore several mobile applications designed for individuals with ASD, specifically those for improving emotional intelligence. Their findings indicated that customizable features, such as the ability to upload familiar faces, greatly enhanced an application's ease of use for children. In addition, they praised applications that included real-world example settings for enabling children to generalize emotion recognition skills beyond the application itself. Therefore, these two settings are both considered in the application design of this study.

2.2 Applications for Emotion Intelligence Training of ASD

A common social challenge for autistic individuals lies in interpreting the actions, facial expressions and words of others and formulating context-appropriate responses (SEN Magazine, 2009). This difficulty is closely tied to emotional intelligence, which encompasses the understanding, identification, and regulation of emotions. As autistic individuals may process emotions differently from neurotypical peers (NHS Transformation Partners, 2017), tailored interventions are crucial. The convergence of technology with therapeutic practice now presents a promising pathway to support personalized emotional intelligence development in ASD.

Several studies have already explored how technology can help autistic people to develop their emotion recognition skills. Madsen et al. (2008) built an application to assist adolescents with ASD in decoding facial expressions during live interactions. Through real-time video analysis via Dynamic Bayesian Networks, it detects facial expressions and provides users with clear visual metrics of the emotional content. These metrics were shown as "emotion bubbles", where the radius of the bubble would resize to reflect the strength of the emotion in the video. A pilot study was conducted with three adolescents with ASD, and the results suggested that the bubbles were perceived as extremely helpful, as the participants attempted to adjust their expressions to make certain bubbles grow or shrink. This application was limited, however, by its adaptability to a real-life conversation, as it is impractical to require a live stream of the face of the speaker in a casual conversation. The study also does not discuss how the skills learned in the exercise could be generalised outside of the use of the application.

Emo-mirror (Pavez et al., 2021) is a recent study which also used artificial intelligence to detect the emotions shown on an individual's face in real-time. In this study, a smart mirror setup was used to display images depicting a particular emotion to the user. The user was able to see their faces next to the example and practice the facial expressions themselves. The application installed in the mirror was trained using a convolutional neural network (CNN), and the image dataset was expanded as users interacted with the mirror. This smart mirror was evaluated by 32 seven to fifteen-year-olds with ASD and then 12 ASD healthcare professionals, and the results suggested that it was accessible and useful. It was also found that the emotion of fear was the hardest emotion for participants to imitate, with neutral being the easiest. Similar to the previous study, there was no discussion of how the application could be generalised to the development of emotion recognition abilities in a real-world context.

A notable issue with the preceding two studies is that they encourage concepts known as masking (Miller et al., 2021) and camouflaging (Summerill and Summers, 2025), where the individual with ASD is encouraged to modify their behaviour and conform to expectations that the "correct" way to behave is how a neurotypical individual would. These applications encourage individuals to modify their facial expressions and mimic expressions to receive positive feedback. However, masking and camouflaging can be hugely detrimental to the development of individuals with ASD, as it can cause increased anxiety, depression, burnout, and exhaustion in autistic individuals (Summerill and Summers, 2025). Therefore, in the application outlined in this paper, a focus was put on enabling a mutual understanding of how all individuals convey emotion, and no aspect of the application would ask any individual to change their expression of emotion.

An avatar refers to a digital representation or icon used to embody a user in virtual spaces, such as video games, or virtual reality platforms (Nowak and Fox , 2018). The Virtual Messenger platform (Fabri et al., 2007) allows autistic people to improve emotional expression and recognition by engaging in avatar-based chats within a virtual learning environment. These avatars were equipped with seven emotional expressions: happiness, surprise, anger, fear, sadness, disgust and neutral. For instance, an avatar might highlight varying degrees of sadness, from slight disappointment to profound sorrow, allowing the user to discern the differences and nuances between these similar emotional states. Their controlled, consistent emotional displays can act as anchors, granting individuals with ASD a sense of predictability and stability in a digital world. An experiment involving 32 participants (ages 21–63, evenly split by gender) was conducted, with individual sessions lasting between 8 and 35 minutes. After practising with expressive avatars in multi-user scenarios, most participants reported high levels of enjoyment, presence, and co-presence. These results suggest that the application holds promise for supporting the development of socio-emotional abilities in adults with ASD.

Using a subset of the facial expression from Virtual Messenger, another structured, three-stage single-user computer system to teach emotion recognition using avatar representations of four basic emotions (happy, sad, angry, frightened) was also developed. The training begins by establishing a recognition baseline, progresses to predicting emotions within simple social scenarios, and culminates in inferring the causes of displayed emotions. A key feature of the system is its exclusive use of schematic avatar faces, rather than real human faces, to portray emotional expressions throughout all stages. An experiment involved 34 children on the autism spectrum (18 with Asperger's, 16 with severe autism), aged 7–16 years. Overall, 30 participants performed significantly better than chance in recognizing avatar emotions, indicating that the avatars are generally effective for emotion recognition. Notably, all four participants who performed at chance level came from the severe autism subgroup, suggesting that avatar-based emotional representations may be less accessible for individuals with greater support needs.

Recent studies (Vasilevska Petrovska and Trajkovski, 2019; Pino et al., 2021) suggest that combining avatars and human faces may enhance emotion recognition in children with ASD, yet the transfer of skills from avatar interaction to real-world human interactions remains unexplored. Therefore, this paper aims to develop an application that addresses this gap and also evaluate its effectiveness.

3 METHODOLOGY

This study proposes EmoBridge: a mobile application designed to bridge the gap in emotion comprehension, starting with identifying emotions in avatars and progressing to understanding human emotions. EmoBridge includes two sections: a gamified section with an adaptive learning algorithm and a browse section with learner customisation considerations.

3.1 The Datasets

The set of images provided within EmoBridge are a subset of three open-source data sets: Facial Expression Research Group 2D Database (FERG-DB) (Aneja et al., 2017), Universitat de les Illes Balears Virtual Facial Expression Dataset (UIB-VFED) (Oliver and Amengual Alcover, 2020) and Real-World Affective Faces (RAF) Database (Li et al., 2017), as shown in Fig. 1 from top to bottom.

FERG-DB consists of over 55,000 facial images across six characters, grouped by seven different emotions. The six characters represent three male and three female faces, with a mix of differing skin tones. The seven emotions which the data set covers are anger, disgust, fear, joy, neutral, sadness, and surprise, which fulfils both basic and intermediate deliverables of the project for the emotions which the application should include. UIBVFED consists of 640 facial expression images across 20 characters which represent several different ethnicities and ages across both male and female genders. This data set covers the same seven emotions as in FERG-DB. Finally, the RAF database contains over 15,000 images of human facial expressions across the same seven emotions, representing all ages, ethnicities, and genders. In this dataset, the emotion 'joy' was labelled 'happiness' instead. To match the avatar emotions with the human emotions, the assumption was made that joy and happiness are the same emotion.

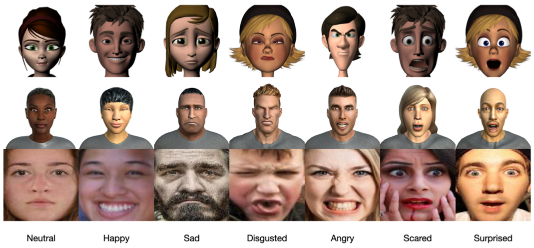

Figure 1: Examples of the expressions within EmoBridge (from top to bottom: FERG-DB, UIBVFED and RAF).

The images in EmoBridge (a total of 480 avatar images and 480 human face images) were selected at random from the above three open datasets. Importantly, the number of each emotion in each dataset was deliberately balanced (see Table 1) to ensure an appropriate representation of each dataset and emotion. It was also ensured that the size of the image dataset used within EmoBridge was sufficient to provide the user with a wide enough breadth of facial expressions, without compromising their experience when running the application. It is worth noting that UIBVFED has more details included in the faces than FERG-DB which makes it more complex and closer to human form. Therefore, our study considered the difficulty to recognize the emotions arising from FERG-DB to UIBVFED, and then RAF.

	FERG-DB	UIBVFED	RAF
Total Count	250	230	480
Female %	50.8%	49.1%	50.6%
Anger	30	30	60
Disgust	40	40	80
Fear	28	20	48
Joy	60	60	120
Neutral	24	20	44
Sadness	40	40	80
Surprise	28	20	48

Table 1: The images composition of EmoBridge

3.2 Gamification-based Section with an Adaptive Learning Algorithm

A gamification-based section has been designed and implemented to facilitate users in developing their emotion recognition skills through the utilisation of the avatar-to-human transition. Individuals with ASD often exhibit differences in attention, including challenges with sustaining attention on demand (Boxhoorn et al., 2018), particularly for non-preferred tasks. Therefore, application design should prioritize brief, engaging segments to align with these attentional patterns. EmoBridge delivers content in sessions of 15 questions each, structured to support brief, focused interactions. This design facilitates engagement with the learning material while incorporating necessary breaks to refocus attention as needed.

As shown in Fig. 2, for each question the user is presented with one facial expression image and asked to identify the emotion being depicted by the avatar or human face. The user has two attempts to successfully identify the emotion from a set of options. On the first attempt, all seven possible emotions are given as options(Fig. 2b). On the second attempt, the number of options is reduced to four emotions (Fig.2d). The three options which are dropped are the incorrect emotion which they chose on their first attempt and two randomly chosen other incorrect emotions.

If the user correctly identifies the emotion on either attempt, the application will reward them with a star (Fig. 2c). If they are incorrect on their first attempt, EmoBridge will display the message 'Have another go!' alongside the reduced set of options (Fig. 2d). If they then incorrectly identify the emotion again, the application will display the message 'Not quite' alongside the correct emotion for the image (Fig. 2e). The decision to employ encouraging phrases, as opposed to terms such as "incorrect" or "wrong," is grounded in evidence that individuals with ASD may exhibit sensitivity to criticism while responding positively to reinforcement (Broadbent and Stokes, 2013). Consequently, the gamified experience is structured to reward correct emotion identification with stars, while offering constructive and encouraging feedback for incorrect responses. Stars earned during gameplay are aggregated and displayed to the user following each 15-question session, as shown in Fig. 2f.

To answer the research question that focuses on the effect of the transition from avatars to human images, an adaptive learning algorithm is proposed and implemented in the gamified session for the selection of images from the 3 datasets mentioned in Table 1 for each question. The adaptive learning algorithm, illustrated in Fig. 3, dynamically selects a dataset for each in-game question based on the user's current emotion recognition skill level, which is inferred from their response history.

As first-time users have no prior response history, the adaptive algorithm does not incorporate user performance data until a threshold of at least five answered questions per dataset has been reached. The system requires the user to complete a minimum of five questions from the basic avatar dataset FERG-DB before advancing to the complex avatar dataset UIBVFED. This same prerequisite applies to UIBVFED dataset before the human dataset RAF is introduced.

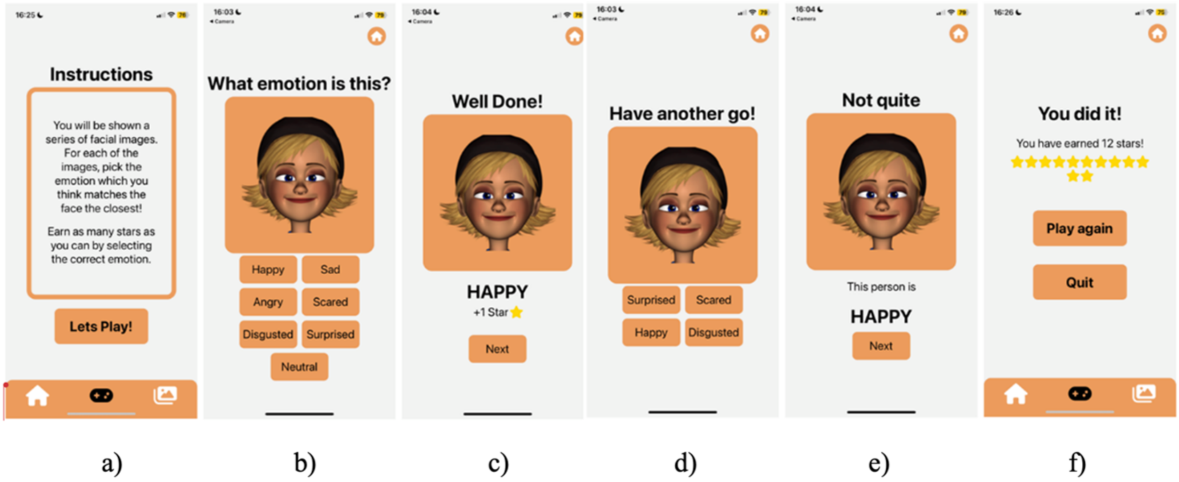

Figure 2: Examples of gamification-based sessions within EmoBridge.

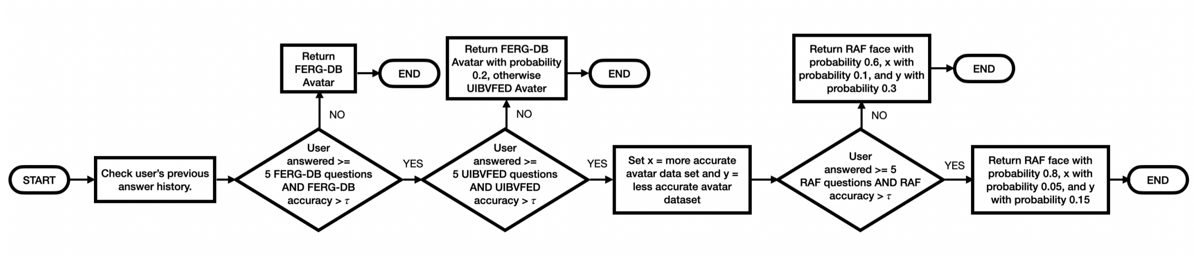

Figure 3: Flowchart of adaptive learning algorithm.

This phased approach ensures that the user's accuracy for each dataset is calculated from a sufficient sample size to guarantee measurement consistency.

As depicted in the flowchart (Fig.3), the algorithm elevates difficulty only when the user's emotion identification accuracy for a dataset exceeds a threshold τ, ensuring mastery before advancement. In the experiment of this study, the adaptive learning algorithm initially employed a 75% threshold for progression. However, the threshold could be manually adjusted downward in EmoBridge in cases where participants with ASD demonstrated considerable frustration as a result of being unable to progress. For each question, the system first determines all datasets suited to the user's current skill level, then selects one at random from this subset. This approach guarantees that every question presents an appropriate, personalized challenge.

Once the user's accuracy meets the threshold for the human dataset, the algorithm concurrently evaluates their performance across the avatar datasets. To ensure continued appropriate challenge, it then adaptively balances the selection, prioritizing questions from the avatar dataset with which the user demonstrates lower accuracy. Because the algorithm operates on a per-question basis, difficulty can decrease if performance on any dataset drops below the threshold. This enables users to experience multiple difficulty levels within one run, ensuring the challenge is continuously personalized to optimize skill development.

3.3 Browse Section with learner customisation considerations

EmoBridge also includes a browse section where the user can view all the images which may appear in the gamification-based section. In this section, the images are categorised by emotion, allowing the user to select a particular emotion of which they wish to see further examples. As shown in Fig. 4, both the avatar images and human images are presented to the user after they choose to study the emotion "Happy". There is no specific order in which the images will be displayed, as a random order has been applied to ensure full integration of the avatar and human emotions. This means that every time the user exits the page for an emotion, and then reopens it, the images may appear in a different order.

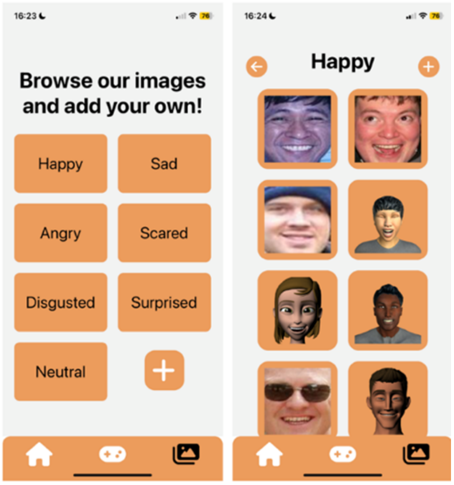

Figure 4: An example of the browse section of EmoBridge. (a) initial menu to select emotion and (b) the screen displaying all images for the emotion happy

Moreover, EmoBridge also allows the user to upload additional images to the application; a feature considering customization. It is intended to be a supplementary tool to aid the development of emotion recognition skills by providing a collection of examples of how emotions can be depicted on a face. Alongside the options for browsing all the emotions, there is also a '+' button in the top-right part of Fig. 4(b) which allows the user to add an image. As shown in Fig. 5, when the '+' button is selected, the application will access the device's camera, enabling users to either capture a new facial photograph or choose an existing one from the local photo gallery. After the user takes their picture, they can add it to the application by pressing 'Save', or they can delete it using the 'Discard' button and take a new one .

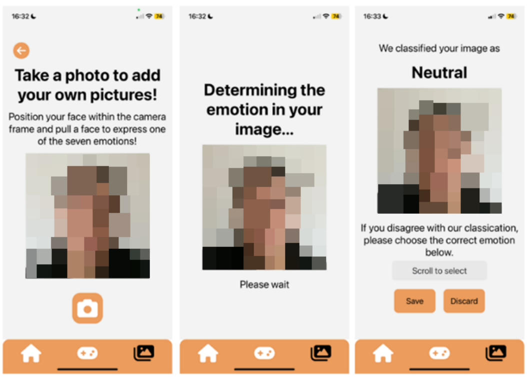

Figure 5: User Interface for uploading users' photos

EmoBridge includes a CNN classification model pre-trained on the 15 thousand images of RAF human face dataset (80% for training and 20% for validation) to classify the uploaded human image. The current model in EmoBridge achieved a validation accuracy of 0.71. To mitigate the impact of erroneous predictions and ensure reliable classification of images, EmoBridge also provides a manual override mechanism, allowing users to rectify incorrect labels if they disagree with the classification. As shown in the right of Fig. 5, the user can overwrite the label provided by the classifier for the image, by clicking "Scroll to select" button to choose from a list of emotions.

4 Experimental Procedure

In order to answer the research question and assess the effectiveness of EmoBridge in aiding children with ASD to develop their emotion recognition abilities, a preliminary experiment was conducted with a group of children aged 7-12, to observe their experience with EmoBridge. Ethics approval was granted by the institutional review board, and written informed consent was obtained from a parent or guardian for each participating child. In total, 11 children who had a formal diagnosis of ASD (7 verbal and 4 non-verbal), and 7 neurotypical children participated in the experiment. The groups of children were age-matched such that the average age of each group was 10 years old.

The experimental procedure is shown in Fig.6. Each child was asked to complete baseline and post-intervention tests to assess their emotion recognition abilities before and after using EmoBridge. To assess the real-world emotion recognition skills, both assessments required the child to identify emotions being depicted in a series of human faces. They were intended to assess the child's abilities to identify each of the seven emotions which are included in EmoBridge. Each test had 20 questions, and each question asked to select which of the seven emotions was represented in the image being shown. The images used in each test were distinct from each other, and from any image used in EmoBridge, to ensure that all images were completely unfamiliar to each child.

After completing the baseline test, each child was asked to interact with EmoBridge for approximately 15 minutes, using both the gamified and browse sections. If parental consent was granted, the child was also encouraged to use the camera functionality to upload their photos to the application. Participants were free to engage with EmoBridge at their own pace. However, to ensure a complete evaluation, participation required interaction with both application sections. Furthermore, a minimum of three rounds in the gamified session was mandatory, continuing until the

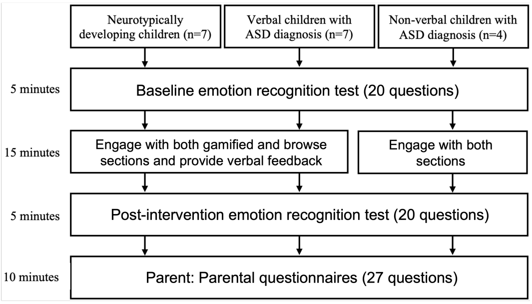

Figure 6: Experimental Procedure

child had engaged with both avatar and human emotion tasks. Additionally, verbal feedback was elicited from the neurotypical and verbal children through open-ended questions regarding user experience, including their general impressions of EmoBridge and specific aspects they liked or disliked. The children were also encouraged to ask any questions they had about EmoBridge, which were then answered and further discussed if appropriate.

Following the children's participation, a parental questionnaire was distributed along with a screen recording of EmoBridge. This recording provided context for the parents, who had not witnessed their child's interaction with EmoBridge during the experimental session. Feedback was obtained from all groups of parents to evaluate parental opinions. In addition to background information, the questionnaire covered opinions on general emotion-recognition applications (5 items), technology acceptance (ease of use and usefulness (Na Nongkhai et al., 2026)) specific to the two sessions of EmoBridge (4 items), user satisfaction (8 items), perceived suitability (4 items), along with 6 open-ended questions.

5 RESULTS

5.1 Participant Engagement

Our experiment was carried out in a one-to-one school setting. Neurotypical children spent an average of 9.57 minutes on the application, whereas children with ASD spent nearly twice as long, at 18.6 minutes.

Unfortunately, the group of non-verbal children with ASD struggled to engage with the baseline and post-intervention tests as they were unable to read and understand the questions being asked. Rather than communicating with words, children with non-verbal ASD prefer to use movements and sounds, and often they use pictures and symbols to communicate. A further two verbal children diagnosed with ASD were also unable to complete the tests. These children were both aged 7. As they were slightly younger than the other participants, their attention span was much shorter, hence they struggled to complete the tasks required for the experiment as they became distracted easily and wanted to do other activities instead. Therefore, although still joining the experiment, for the analysis of baseline and post-intervention test results, these two young children and all the four non-verbal ASD children are excluded, as no results could be obtained.

As shown in Fig.7, for both participant groups, the post-intervention tests revealed a significant improvement in emotion recognition accuracy compared to their baseline performance, indicating a positive effect of the application. Notably, children with ASD exhibited a greater average score increase (+3.80, +19%) than their neurotypical peers (+2.29, +11.5%). Furthermore, the intervention appears to have had a particularly meaningful impact on the children with ASD. While this group exhibited a wide variance in baseline ability ($S.D. = 5.26$), reflecting the heterogeneous nature of the spectrum, their performance became markedly more consistent after using the application ($S.D. = 3.24$). In contrast, the neurotypical children demonstrated high consistency across both tests ($S.D. \approx 1.4$). This convergence in performance, alongside the overall score increase, provides strong evidence that EmoBridge effectively supports emotion recognition skill development, especially for verbal children on the autism spectrum. Another notable finding was that after using EmoBridge, verbal children with ASD achieved emotional recognition performance (Mean=15.00) comparable to that of neurotypical children prior to any intervention (Mean=15.71). Despite the limited sample size, these results (encompassing score increases, performance convergence, and reduced score dispersion) provide converging evidence that EmoBridge offers substantial support for emotion recognition development, particularly for verbal children on the autism spectrum.

Each child was observed while using EmoBridge, and any verbal feedback they provided was recorded. Additionally, the children were asked questions about the application if they appeared willing to provide answers. Among them, ten children (56%) added an image to the application, and two (11%) chose to add multiple images. When asked what they liked the best about the application, nine children reported the process of uploading an image of their face as their favourite part. Furthermore, all children who used the camera to upload an image displayed enthusiasm for the task. They expressed satisfaction upon viewing their image within the gallery alongside others. How-

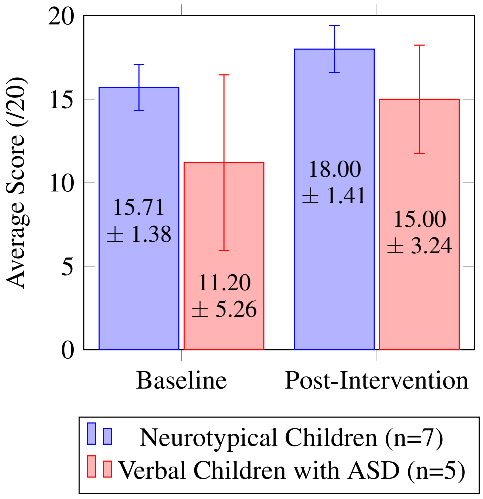

Figure 7: Baseline vs Post-Intervention Results.

ever, one child reported that although they enjoyed this function, they did not like that the classifier incorrectly labelled their disgusted face as being scared. They indicated that their enjoyment of the application would likely increase if the classifier had labelled their image correctly.

For two participants with ASD whose response accuracy on questions involving the avatar dataset UIBVFED could not reach the initial 75% threshold of the algorithm, the criterion was manually lowered. This intervention was prompted by observed frustration and the goal of enabling them to access the human image dataset. This finding points to the potential need for personalized threshold calibration within gamified session for some children with ASD. Notably, for both children, response accuracy was higher for questions with RAF human images than with UIBVFED avatars. On the other hand, several of the children (2 neurotypical and 3 with ASD diagnosis) made comments about finding some of the avatars from the UIBVFED dataset a bit weird, and those with ASD were also observed to struggle more to recognise emotions from these avatars. Both results suggest some alternative, complex avatar dataset may need to be found to replace the UIBVFED dataset in future studies.

5.2 Parental Questionnaire

A parental questionnaire, accompanied by a screen recording of EmoBridge, was distributed to all parents who permitted their children to participate in the study. The questionnaire aimed to assess the application's concept, usability, and design suitability. In total, 10 out of 18 parents responded. The demographics of the respondents included parents of 2 neurotypical children and 8 children diagnosed with ASD. Within the ASD group, the data was further split into parents of verbal children ($n = 5$) and non-verbal children ($n = 3$), spanning an age range from 7 to 11 years.

To establish a baseline and assess the suitability of the application, parents were asked to rate the extent to which they believe their child struggles with emotion recognition on a scale of 1 (Not at all) to 5 (Very much). Both parents of neurotypical children rated the struggle as 1, whereas the average rating for parents of children with non-verbal ASD was 4.0 (S.D. = 1.0) and verbal ASD was 2.6 (S.D. = 0.89), confirming the target demographic's need for intervention.

All 10 parents agreed that an application to aid emotion recognition would be useful in general. When asked about their own child specifically, 90% of parents agreed or strongly agreed such an application would be beneficial and would encourage their child to use it if available. Notably, the only parent who expressed neutrality was the parent of a neurotypical child who had previously indicated their child had no struggle with emotions.

Parents rated the ease of use and the perceived usefulness of the gamified and browse sections of EmoBridge on a 5-point scale. Parents of both neurotypical children and verbal ASD children agreed that EmoBridge is easy to use (Fig. 8). In contrast, the parents of non-verbal ASD children have a more neutral attitude. This finding was consistent with their qualitative feedback, where they noted that "for non-literate children, the multiple options to be read could create barriers to engagement." This parent suggested integrating Picture Exchange Communication System (PECS) cards (Bondy and Frost, 1994) alongside text labels, slowly fading them out as accuracy improves. This insight highlights that whilst EmoBridge is highly usable for literate children, future iterations should reduce the reliance on text to accommodate non-literate users effectively.

Regarding the perceived usefulness (Fig. 9) of EmoBridge, results were generally positive but highlighted specific demographic nuances. Across both sessions, parents of verbal children with ASD reported a slightly higher average usefulness rating than parents of non-verbal children with ASD, with both groups showing low standard deviations (SD). Parents of neurotypical children provided the lowest average rating, though with substantially greater variability (SD $>$ 2). Specifically, the lowest rating came from the neurotypical parent who reported their child had no issues with recognising emotions, suggesting that they simply did not perceive a need for the intervention, rather than a flaw in the application design itself.

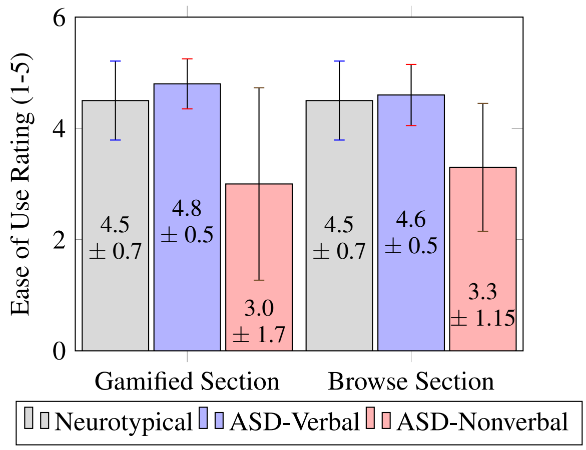

Figure 8: Ease of use ratings grouped by section.

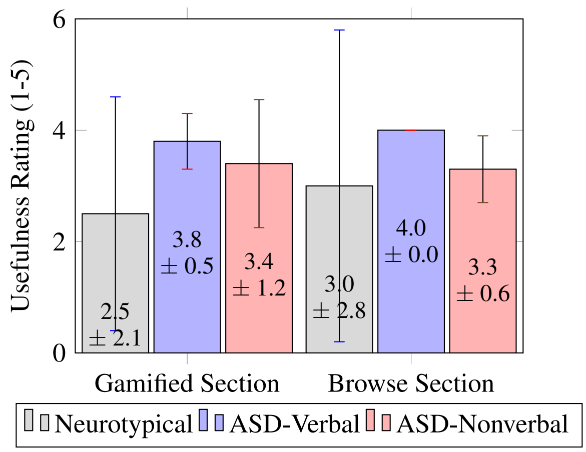

Figure 9: Perceived usefulness ratings grouped by section.

The key features of EmoBridge, specifically the inclusion of user-uploaded photos and the transition from avatars to human images, were evaluated by the parents. All of parents agreed that the ability to add familiar faces, such as family members, would help their child develop emotion recognition skills. One parent specifically commented that their child "really enjoyed taking her own facial expression picture. Regarding the avatar-to-human transition, 80% of parents believed this mechanism would be effective, and 90% agreed that the general inclusion of both avatar and human imagery was beneficial. However, the design of the avatars generated mixed feedback. Whilst the first avatar dataset (FERG-DB) was generally well-received, the second dataset (UIBVFED) was considered less appropriate, with 50% of parents rating its suitability as 3 out of 5 or lower. One parent explicitly noted that some avatars were "tricky to read," suggesting that stylised avatars may introduce unnecessary ambiguity for some children with ASD.

In terms of general perception, results indicated a predominantly positive response. When asked if their child would enjoy the application, 80% of participants agreed or strongly agreed, and 90% of parents confirming they would download the application if it were available with 50% of respondents anticipating using the app at least a few times a week, 20% a few times a month, and 10% each for once a month, once a week, and never. It should be noted that the parent who rated "never" is the same parent of the neurotypical child who has no difficulty in emotion recognition. Finally, parents were asked to rate, on a scale of 1 to 5, the specific question: "In general, to what extent do you believe that the application will help your child to develop their emotion recognition abilities?" Parents of children with Verbal ASD reported the highest level of satisfaction, with a mean rating of 4.2 (SD=0.45), indicating that the application effectively meets the needs of this core demographic. In contrast, both the neurotypical and nonverbal ASD groups reported a lower mean rating of 3.0 (SD = 2.82) and 3.3 (SD=1.15) respectively.

Finally, the parents had the opportunity to provide any further comments on the application. Two parents said they had observed that their child's emotion recognition abilities had improved to some degree in the time after they had used EmoBridge. Four of the parents with ASD children left comments commending the simplicity and ease of use of EmoBridge, with one parent highlighting the importance of this type of application for individuals with ASD. Moreover, constructive feedback from a parent of a neurotypical child suggested adding voice-overs for emotion words to assist younger readers, reinforcing the need for multi-sensory accessibility features.

6 CONCLUSION, LIMITATION AND FUTURE WORK

In this research, a mobile application EmoBridge has been developed to support individuals with ASD in developing emotional recognition skills. For the first time, an application was proposed which utilises a transition from avatars to human faces by integrating images from three open-source facial expression datasets, with the intention that this transition may help individuals with ASD to successfully identify emotions. It incorporated an adaptive learning algorithm to determine the speed of the transition based on the user's emotion recognition skill level and a classification model to categorise user-uploaded images.

The experimental results obtained provide positive evidence that EmoBridge does support children aged 7-12 to develop their emotion recognition skills.

Excluding non-completers (the non-verbal ASD cohort and the two youngest verbal ASD participants), analysis of the remaining children revealed a post-intervention score gain. The average improvement was significantly larger in the verbal ASD group than in the neurotypical group. Additionally, a convergence in performance was observed within the verbal ASD group. Moreover, following the intervention, verbal children with ASD reached an emotional recognition level that approached the pre-intervention baseline of neurotypical children. Therefore, it is suggested that EmoBridge did support verbal ASD children to develop their emotion recognition skills.

Furthermore, results from the parental questionnaire indicated that most of the parents agreed that EmoBridge is easy for their child to use and useful for their child to develop emotion recognition abilities. Notably, except for one parent of a neurotypical child who had no problems with recognising emotion, the rest expressed a willingness to use EmoBridge if available.Whilst the majority considered the application suitable, feedback from the parent of a non-verbal child highlighted the need for enhanced accessibility. This respondent provided suggestions regarding the adaptation of the interface for non-literate users, specifically the integration of symbol-based communication to support non-verbal children.

In summary, the approach of introducing avatars and then transitioning to real human faces helps bridge the gap between virtual and real-world interactions, allowing individuals with ASD to adapt gradually to more complex emotional cues. The gamified section of EmoBridge provides positive reinforcement, boosting user confidence and motivation, which is particularly important for individuals with ASD. Furthermore, the browse section encourages the user to upload their own photos. This customization serves a dual purpose: it increases engagement and boosts motivation during the gamified session by incorporating familiar photos into gameplay.

However, this study has several limitations. Firstly, due to the small sample size, the conclusions which have been drawn cannot be reliably generalised to the wider population of children with ASD. Given that the average age of diagnosis of ASD is 14.5 years (Russell et al., 2021), the younger children who have diagnoses of ASD generally experience more extreme symptoms, therefore, it was difficult to collect a large sample of participants who had a diagnosis and were able to complete the tasks required for the study. Nevertheless, further investigation with a larger sample is needed to determine the application's ability to transfer the emotion recognition skill to real-world interactions in the long term. Each participant should interact with EmoBridge multiple times on various occasions, and their emotion recognition skills should be monitored continuously to assess any development, to more precisely answer the research question if our application can support individuals with ASD to develop transferable emotion recognition abilities by utilising a transition from avatar faces to human faces.

Secondly, the age range of the participants needs to be carefully reconsidered because the mental load required for younger children to complete a series of tests and tasks is much higher than for older children. Hence the concern with recruiting younger children to participate in the further study surrounded their ability to complete all the tasks accurately to provide adequate results for conclusions to be drawn. Additionally, individuals with ASD tend to have a reduced attention span,This is supported by this study that two of 7 year old verbal ASD children were unable to maintain focus and complete the tests.

Thirdly, the results suggest that non-verbal children with ASD struggled to interact with EmoBridge, and feedback from parents revealed that this was because EmoBridge used words to describe the emotions being represented. This suggests that EmoBridge is currently limited in its accessibility and helpfulness for the non-verbal ASD children who potentially need it most. Future work should consider adding more symbols and pictures, such as PECS card images which are already familiar to non-verbal ASD, to go with the words, and then slowly reducing them as the user's response accuracy increases. Also, the non-verbal group of children all struggled to interact with the baseline and post-intervention tests, hence the test content for them needs to be amended accordingly using PECS card images.

Finally, whilst the adaptive learning algorithm in EmoBridge currently determines the images from the datasets based on user accuracy, it does not affect the frequency of the selection of emotions presented in the gamified section. In future, we plan to adapt the algorithm to focus on the emotions which the user finds more challenging to identify.

ACKNOWLEDGEMENTS

This work was supported in part by Horizon Europe MSCA-2022-SE (Project UNITE, Grant 101129618).

REFERENCES

Aburukba, R., Aloul, F., Mahmoud, A., Kamili, K., and Ajmal, S. (2017). Autiaid: A learning mobile application for autistic children. In *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE.

Alves, S., Marques, A., Queirós, C., and Orvalho, V. (2013). Lifeisgame prototype: A serious game about emotions for children with autism spectrum disorders. *PsychNology Journal*, 11:191–211.

Aneja, D., Colburn, A., Faigin, G., Shapiro, L., and Mones, B. (2017). Modeling stylized character expressions via deep learning. In Lai, S.-H., Lepetit, V., Nishino, K., and Sato, Y., editors, *Computer Vision – ACCV 2016*, volume 10112 of *Lecture Notes in Computer Science*, pages 136–153, Cham. Springer.

Baron-Cohen, S., Golan, O., and Ashim, E. (2009). Can emotion recognition be taught to children with autism spectrum conditions? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3567–3574.

Bondy, A. S. and Frost, L. A. (1994). The picture exchange communication system. *Focus on Autistic Behavior*, 9(3):1–19.

Boxhoorn, S., Lopez, E., Schmidt, C., Schulze, D., Hänig, S., and Freitag, C. M. (2018). Attention profiles in autism spectrum disorder and subtypes of attention-deficit/hyperactivity disorder. *European Child amp; Adolescent Psychiatry*, 27(11):1433–1447.

Broadbent, J. and Stokes, M. A. (2013). Removal of negative feedback enhances wcst performance for individuals with asd. *Research in Autism Spectrum Disorders*, 7(6):785–792.

Fabri, M., Elzouki, S. Y. A., and Moore, D. (2007). *Emotionally Expressive Avatars for Chatting, Learning and Therapeutic Intervention*, page 275–285. Springer Berlin Heidelberg.

Li, S., Deng, W., and Du, J. (2017). Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2852–2861.

Madsen, M., el Kaliouby, R., Goodwin, M., and Picard, R. (2008). Technology for just-in-time in-situ learning of facial affect for persons diagnosed with an autism spectrum disorder. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 19–26. ACM.

Miller, D., Rees, J., and Pearson, A. (2021). “Masking Is Life”: Experiences of masking in autistic and nonautistic adults. *Autism in Adulthood*, 3(4):330–338.

Mohammad, H. and Abu-Amara, F. (2019). A mobile social and communication tool for autism. *International Journal of Emerging Technologies in Learning (iJET)*, 14(19):159.

Na Nongkhai, L., Wang, J., Wynn, A., and Mendori, T. (2026). Evaluating adaptive and generative ai-based feedback and recommendations in a knowledge-graph-integrated programming learning system. *Computers and Education: Artificial Intelligence*, page 100526.

National Autistic Society (2022). What is autism? https://www.autism.org.uk/advice-and-guidance/what-is-autism. Accessed: 22 July 2023.

NHS Transformation Partners (2017). Children and young people with autism spectrum disorder. https://www.transformationpartners.nhs.uk/wp-content/uploads/2018/11/Children-and-young-people-with-austism-spectrum-disorder-case-for-change-Oct-2017.pdf. [Accessed 04-01-2026].

Novack, M. N., Hong, E., Dixon, D. R., and Granpeesheh, D. (2018). An evaluation of a mobile application designed to teach receptive language skills to children with autism spectrum disorder. *Behavior Analysis in Practice*, 12(1):66–77.

Nowak, K. L. and Fox , J. (2018). Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations on communication. *Review of Communication Research*, 6:30–53.

Oliver, M. M. and Amengual Alcover, E. (2020). Uibvfed: Virtual facial expression dataset. *PLOS ONE*, 15(4):e0231266.

Papoutsi, C., Drigas, A. S., and Skianis, C. (2018). Mobile applications to improve emotional intelligence in autism – a review. *International Journal of Interactive Mobile Technologies (iJIM)*, 12(6):47.

Pavez, R. et al. (2021). Emo-mirror: A proposal to support emotion recognition in children with autism spectrum disorders. *Neural Computing and Applications*, pages 1–12.

Pino, M. C., Vagnetti, R., Valenti, M., and Mazza, M. (2021). Comparing virtual vs real faces expressing emotions in children with autism: An eye-tracking study. *Education and Information Technologies*, 26(5):5717–5732.

Russell, G., Stapley, S., Newlove-Delgado, T., Salmon, A., White, R., Warren, F., Pearson, A., and Ford, T. (2021). Time trends in autism diagnosis over 20 years: a uk population-based cohort study. *Journal of Child Psychology and Psychiatry*, 63(6):674–682.

SEN Magazine (2009). How does autism affect behaviour? https://senmagazine.co.uk/content/specific-needs/autism-asd/177/how-does-autism-affect-behaviour/. Accessed: 14 April 2022.

Summerill, J. and Summers, S. A. (2025). The consequences of social camouflaging in autistic adults: A systematic review. *Research in Autism*, 121-122:202556.

Vasilevska Petrovska, I. and Trajkovski, V. (2019). Effects of a computer-based intervention on emotion understanding in children with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 49:4244–4255.

Yee, H. S. S. (2012). Mobile technology for children with autism spectrum disorder: Major trends and issues. In *2012 IEEE Symposium on E-Learning, E-Management and E-Services*, page 1–5. IEEE.

Zeidan, J. et al. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, 15(5):778–790.