Comparing virtual vs real faces expressing emotions in children with autism: An eye-tracking study

Maria Chiara Pino[1,2] · Roberto Vagnetti[1] · Marco Valenti[1,2] · Monica Mazza[1,2]

Abstract

Difficulties in processing emotional facial expressions is considered a central characteristic of children with autism spectrum condition (ASC). In addition, there is a growing interest in the use of virtual avatars capable of expressing emotions as an intervention aimed at improving the social skills of these individuals. One potential use of avatars is that they could enhance facial recognition and guide attention. However, this aspect needs further investigation. The aim of our study is to assess differences in eye gaze processes in children with ASC when they see avatar faces expressing emotions compared to real faces. Eye-tracking methodology was used to compare the performance of children with ASC between avatar and real faces. A repeated-measures general linear model was adopted to understand which characteristics of the stimuli could influence the stimuli's fixation times. Survival analysis was performed to understand differences in exploration behaviour between avatar and real faces. Differences between emotion recognition accuracy and the number of fixations were evaluated through a paired t-test. Our results confirm that children with autism have higher capacities to process and recognize emotions when these are presented by avatar faces. Children with autism are more attracted to the mouth or the eyes depending on the stimulus type (avatar or real) and the emotion expressed by the stimulus. Also, they are more attracted to avatar faces expressing negative emotions (anger and sadness), and to real faces expressing surprise. Differences were not found regarding happiness. Finally, they show a higher degree of exploration of avatar faces. All these elements, such as interest in the avatar and reduced attention to the eyes, can offer important elements in planning an efficient intervention.

Keywords Autism spectrum condition (ASC) · Facial emotional expression · Virtual faces · Avatar · Eye tracking

✉ Roberto Vagnetti
    roberto.vagnetti.rv@gmail.com

Extended author information available on the last page of the article

1 Introduction

Individuals with autism spectrum condition (ASC) are characterized by difficulty in recognizing and responding to emotions conveyed by the face; their main difficulty is responding appropriately to the emotions of other people (Cassidy et al., 2016). Although the literature has often presented mixed data regarding this deficit, a meta-analysis including 48 studies showed that in subjects with ASC this deficit is indeed present (Uljarevic & Hamilton, 2013); moreover the meta-analysis showed that individuals with ASC have difficulty in recognizing all primary emotions except happiness, albeit marginally. Most eye-tracking studies have investigated emotion recognition in individuals with ASC using real faces and have provided important insights regarding their processing of faces. We know that subjects with ASC placed in front of emotion-expressing faces show altered eye patterns compared to controls with little interest in social stimuli such as eyes (Reisinger et al., 2020; Tsang, 2018). Indeed, attention to social stimuli is considered an early diagnostic biomarker in childhood (Elsabbagh et al., 2014; Jones et al., 2016). Some studies have suggested different visual patterns in the exploration of faces depending on the emotion presented, with one study finding that children with ASC, similarly to controls, look at their eyes for longer when viewing negative emotions (de Wit et al., 2008). Another study (Tsang, 2018) suggested that individuals with ASC use a rule-bound categorical thinking approach to understand facial emotions through a categorical rule (e.g. happiness is represented by a curled-up smile), suggesting that processing style may also influence the understanding of emotions. Moreover, children with ASC could experience more pronounced difficulties when faces express complex emotions (Tsang, 2018). Also, it is worth mentioning that for a correct recognition of an emotion, individuals dwell on specific parts of the face, mainly the mouth or the eye area, depending on the emotion being expressed by the face (Wegrzyn et al., 2017). One study also found an influence of the gender of the person expressing the emotion (Lambrecht et al., 2014). These studies show how many variables can affect recognition and visual patterns when participants are asked to recognize the emotion on the presented face.

One of the main limitations of rehabilitation treatment based on emotions directed at individuals with ASC is that face-to-face or group interventions may not be appropriate for this clinical group. ASC individuals are not very attracted to the human face, perhaps because it is too complex and difficult to interpret for them. Amaral and collaborators (Amaral et al., 2018) suggest that the emotional recognition difficulty in subjects with ASC depends on deficits in the interpretation of others' intentions from gaze direction or other social attention cues. Infants and young children with autism demonstrate impairments in both initiating and responding to joint attention bids in naturalistic settings (Caruana et al., 2018). For this reason, many rehabilitative interventions use technology such as robots, avatars and virtual environments to teach new skills to children with ASC (Orvalho et al., 2009; Jarrold et al., 2013; Bekele et al., 2014; Aresti-Bartolome & Garcia-Zapirain, 2015; Lahiri et al., 2015; Newbutt et al., 2016; Liu et al., 2017; Elgarf et al., 2017; Shoaib et al., 2017; Papathomas & Goldschmidt, 2017; Ravindran et al., 2019; Yuan & Ip, 2018;

Khowaja et al., 2019; Valencia et al., 2019; Rojo et al., 2019; Herrero & Lorenzo, 2020; Di Mascio et al., 2020). Soares et al. (2021) conducted a meta-analysis comparing randomized controlled trials of face-to-face and technological interventions for children and adolescents with autism to improve social skills. These interventions included the use of computer-based software, computer avatars and therapeutic robots and found that both types of intervention resulted in comparable improvements in social skills. On the basis of this result, the authors suggest the implementation of a combination of the two intervention types as a future perspective. Technological interventions that include the use of 3D characters are often used in the context of autism to improve social skills (Kandalaft et al., 2013; Yuan & Ip, 2018), where not only virtual avatars but also complex virtual scenarios or situations that can simulate a real-world situation may be present (Didehbani et al., 2016). The literature concerning the study of avatar faces is still limited, but the evidence suggests that they can be used for intervention purposes. Hopkins and collaborators (Hopkins et al., 2011) assessed the efficacy of FaceSay, a computer-based social skills training program for children with ASC showing improvements in emotion recognition and social interactions. FaceSay allows participants to practise attending to eye gaze, discriminating facial expressions, and recognizing faces and emotions in FaceSay's structured environment with interactive, realistic avatar assistants. Moreover, Kumazaki et al. (2019) demonstrated that children with autism responded more often to a simple humanoid robot and simple avatar than they did to a human. Thus, social skills training, performed in the past by human operators, has been automated through interaction with an avatar (Tanaka et al., 2017). Forbes et al. (2016) demonstrated the potentiality of using an avatar to induce facial mimicry and to improve the ability to recognize and produce emotional facial expressions. Azevedo et al. (2018) used avatars in three different tasks, including an emotion recognition task, on five ASC children where they described improvements in emotion recognition between multiple sessions. Santos et al. (2019) proposed the development of a serious game, also aimed at children with autism, that involves the use of a virtual avatar (called ZECA) expressing emotions in order to facilitate communication and understanding of emotions in the child. A study compared a group of adolescents with ASC with a control group who were asked to identify facial expressions shown by avatars (Bekele et al., 2014). In this study, no differences were found in terms of accuracy of emotion recognition between the two groups, while differences emerged with regard to fixation patterns, where TD children fixed more on the mouth area. The authors also point out that virtual reality could potentially guide and alter gaze processing and attention to enhance facial recognition.

Currently, the literature agrees on the utilization of avatars for harnessing and potentially increasing motivation and interest in social stimuli (such as people) in individuals with ASC.

Bekele et al. (2014) support the hypothesis that deficits in emotion and face recognition for individuals with ASC are related to fundamental differences in information processing. Grossard et al. (2018) affirmed that facial emotional expression is a complex developmental process influenced by several factors that need to be acknowledged in future research. Contrary to what has been said so far, Carter et al. (2014) suggested that even avatars that provide live, responsive interactions are not

superior to human therapists in eliciting verbal and non-verbal communication from children with autism.

Interventions with virtual reality and their applications in patients with ASC are of considerable interest; in fact, a strong point of these technologies is that they make it possible to simulate various situations in a controlled environment. However, to the best of our knowledge, no studies have compared the two types of stimuli. Thus, it is necessary to understand what the differences are in the use of virtual and real situations, starting with the simplest and most substantial difference: the virtual stimulus compared to the real one.

ASCs are characterized by a deficit in emotion recognition (Uljarevic & Hamilton, 2013), and a study found that ASC individuals and healthy controls perform at the same level when it comes to emotion recognition when observing avatar faces expressing emotions (Bekele et al., 2014). Our hypothesis is that avatar faces expressing emotion could facilitate emotion recognition better than real faces. According to this hypothesis, we expect children with ASC to be more interested in avatar faces than real faces, thereby showing a longer duration of fixations and enhanced exploration of the stimulus. Furthermore, we investigate which of the avatar face elements in emotion expression could attract the attention of children with autism more than human faces.

2 Method

2.1 Participants

Twenty-nine children with ASC were selected by the Reference Regional Centre for Autism (anonymized, age range 5–11 years) and participated in the study. The ASC diagnosis was provided by an experienced clinician and a senior psychiatrist, according to the criteria of the DSM-5 (APA, 2013), while a senior neuropsychologist conducted the neurocognitive evaluation, and a research assistant and a doctoral student in methodology constructed and performed the experimental task. The ASC diagnosis of patients was confirmed using the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 2012). The VMA was assessed with the Test for Reception of Grammar – Version 2 (TROG-2; Bishop, 2003; Suraniti et al., 2009). For details see Table 1.

Informed consent was given, and accepted, by all the children's parents before the study. Ethical approval was obtained from the hospital's Ethics Committee

Table 1 Demographical and clinical data of ASC participants	Mean (s.d.)
Chronological age (in years)	8.77 (1.42)
Verbal mental age (in years)	12.54 (4.15)
Gender	25 M; 4 F
ADOS-Social communication and social interaction	11.71 (3.25)
ADOS-Repetitive and stereotyped behaviour	2. 71 (0.75)
ADOS total scores	14.43 (3.35)

(anonymized). The Ethics Committee approved the experimental protocol, number 186061/17, before the recruitment of the participants. The exclusion criteria for recruitment were: a history of neurological disease, including epilepsy, head trauma, mental retardation or pharmacological treatment.

2.2 Emotional recognition task: Real versus avatar

We know that children with ASC experience a delay in the development of emotion recognition compared to controls rather than an absence of this capacity (Pino et al., 2017), so we used basic emotions to facilitate emotion recognition and to assess the effect due to avatar or real faces and not compromise their performance with more complex emotions. Moreover, we utilized the emotions suggested by Franco et al. (2014) based on terms known by children. Thus, the set of stimuli was balanced by the number of positive and negative emotions.

The task consisted of 16 stimuli divided into two stimulus types, with eight representing real faces and the other eight representing avatar faces. Each stimulus type showed four emotions (happiness, surprise, sadness and anger), which were presented by a male and a female face to counterbalance any stimuli or gender bias. Each stimulus was presented for seven seconds.

Each child was asked to observe the image and identify the emotion corresponding to the presented stimulus; therefore the number of correct answers was considered for the analyses as a measure of accuracy in emotion recognition where the score range was 0–8 for both the avatar and real emotion recognition scores. Real face stimuli were taken from the Karolinska Directed Emotional Faces database (Lundqvist et al., 1998), while avatar faces were created through FaceGen (Singular Inversions Inc, 2009). Two professional psychologists were then asked what emotion the avatar represented, showing a high degree of agreement (Cohen's kappa = .86). Furthermore, they were presented to a sample of 20 typically developing children with the same age range as our participants, showing a mean accuracy on emotion recognition of above 70%. An example of the avatar stimuli is presented in Fig. 1.

The task was performed using Tobii T120 Eye Tracker equipment consisting of a GL-2760-LED backlit monitor with a resolution of 1920×1080 pixels, which both presented the stimuli and recorded gaze. This eye-tracking system is non-invasive and the subject has little indication that eye movements are being tracked; artificially constraining head movements is not required. The system tracks both eyes to an accuracy of 0.5 degrees at a sampling rate of 60 Hz. The Tobii equipment was connected to a Lenovo laptop computer (Windows 7 Professional) that was used to run the tasks. Calibration procedures, stimulus creation, data acquisition and visualization were performed using the Tobii Studio™ Analysis Software.

2.3 Procedure

All children were tested once, in a quiet, darkened room. The experiment started with a calibration phase that was immediately followed by the test phase. During calibration, a cartoon was presented in the centre of the screen. When the children

Fig. 1 Example of an avatar stimulus expressing anger

started to look at the smiley face it moved to the top left corner of the screen and remained in this position until the toddler fixated on it. Then it moved to the bottom right corner and remained in this position. These three positions were used to compute the pupil–corneal reflection from three points on the screen, allowing the system to derive gaze direction during test phases. The calibration accuracy was checked and the calibration procedure was repeated if necessary. After the calibration phase, the emotion recognition task was administered to the participant. Rectangular areas of interest (AOIs) were defined manually for each image in the displays. Two AOIs were created, namely the eyes and the mouth.

At the end of the testing session, the participant received a reward (i.e. coloured stickers). Two gaze parameters were analysed for each stimulus: (1) total fixation duration (TFD) – the sum of the duration of all fixations within the AOI during the presentation of a given display, thus indicating how much attention had been given to the stimulus; and (2) the number of fixations (NF), where a fixation event was defined by the Tobii fixation filter (I-IV filter) as any occasion on which the direction of gaze remained within 0.5 degrees of the visual angle for at least 100 ms, informing us about children's exploratory behaviour. Data for all AOIs were normalized concerning the total area of the image.

2.4 Statistical analysis

A repeated-measures t-test was performed between the avatar emotion recognition score and the real emotion recognition score to verify differences in emotion recognition accuracy between avatar and real stimuli.

A repeated-measures GLM was performed to assess differences in participants' TFD between the two stimulus types (real and avatar) for stimuli gender (male and female) emotion (happiness, surprise, sadness and anger) and AOIs (eye and mouth), with any significant interaction then being further explored through a post hoc test. Specifically, through GLM we wanted to understand whether any stimulus characteristics could influence time spent on AOIs.

A repeated t-test was performed to compare the number of fixations between real and avatar stimuli in order to compare exploration behaviour.

Furthermore, a survival analysis was performed to understand whether, after a certain number of fixations, participants tended to interrupt the exploration earlier depending on stimulus type. Survival function was described through the Kaplan–Meier method, which gives the probability that an individual will "survive", or in our case continue exploring, after a particular time t, where we consider t as the number of fixations.

For the analysis, alpha was set at .05, and multiple comparisons after GLM were corrected through the Bonferroni method. Analyses were performed using R (R Core Team, 2020) and the *survival* package (Therneau, 2020).

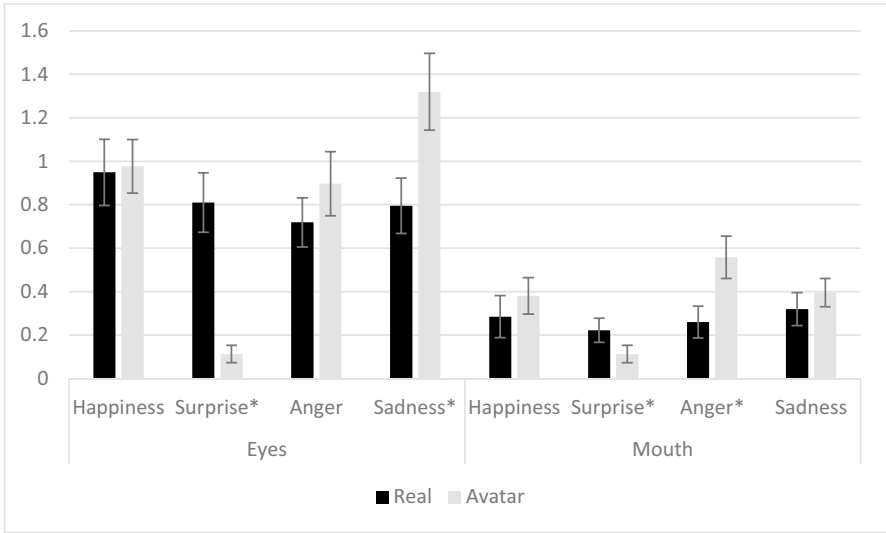# 3 Results

3.1 Emotion recognition

We find a significant difference between the avatar emotion recognition score and the real emotion recognition score (t (28) = 2.36; p = .02), where the accuracy for avatar stimuli (M = 5.42, SD = 1.07) was greater than that for real stimuli (M = 4.84, SD = 1.01), showing a medium effect size (Cohen's d = .54).

3.2 Eye tracker data – Total fixation duration

A significant interaction was found between stimulus type, stimulus gender and AOI ($F_{1, 28}$ = 7.68, p = .01, η_p^2 = .21), as well as between stimulus type, emotion and AOI ($F_{3, 84}$ = 16.5, p < .001, η_p^2 = .37) and between stimulus type and emotion ($F_{3, 84}$ = 21.3, p < .001, η_p^2 = .43).As regards the stimulus type, stimulus gender and AOI interaction, from the post hoc test we did not find any significant differences after the Bonferroni correction that could explain the interaction.

With regard to the interaction between stimulus type, emotion and AOI, we found that TFD on real faces was higher on eyes (M_{dif} = .810, SE = .137, p < .001) and the mouth (M_{dif} = .222, SE = .056, p < .001) when the emotion was surprise; by contrast, TFD was higher for avatar faces on eyes when the emotion was sadness (M_{dif} = −.525, SE = .107, p < .001) and on the mouth when the emotion was anger (M_{dif} = −.299, SE = .104, p = .007).

Finally, the interaction between stimulus type and emotion revealed higher TFD for real faces when the emotion was surprise (M_{dif} = .566, SE = .090, p < .001); by contrast, we found higher TFD for avatar faces when the emotions were anger

*p< .05

Fig. 2 Differences between real and avatar faces of TFD on AOIs during emotion visualization. *$p < .05$

($M_{dif} = -.206$, SE $= .080$, $p < .001$) and sadness ($M_{dif} = -.313$, SE $= .065$, $p < .001$). Results on TFD are reported in Figs. 2 and 3.
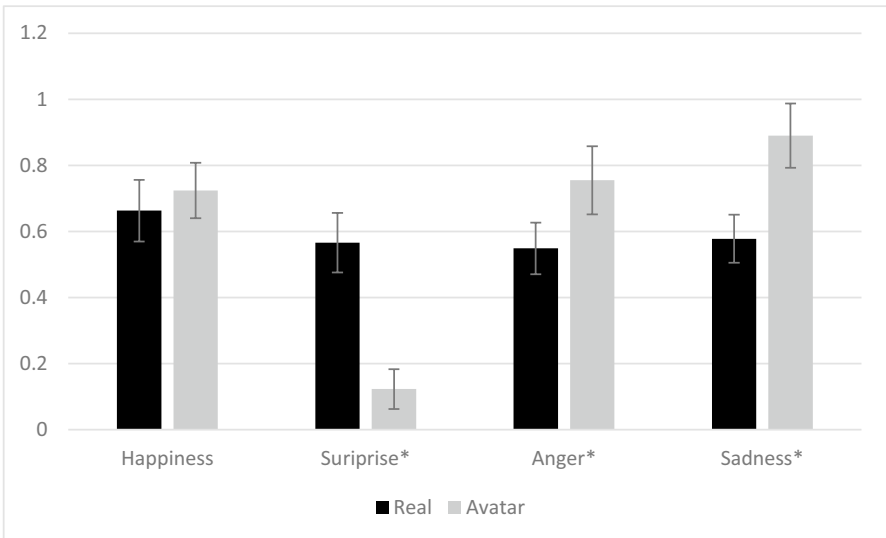
* p< .05

Fig. 3 Differences on emotions' TFD between real and avatar faces. * $p < .05$

3.3 Eye tracker data – Number of fixations

The t-test revealed a significant difference in the number of fixations ($t_{28} = 4.79$, $p < .001$), which was higher for avatar faces (M = 41.5, SD = 25.3) than for real faces (M = 30.3, SD = 20.6), with a large effect size (Cohen's d = .91).

From the survival analysis, we can see that a slightly larger proportion of participants tended to continue exploring avatar faces as the number of fixations increased, although the two curves only tend to be significantly different ($\chi^2(1) = 3.8$, $p = .05$). Survival curves are reported in Fig. 4.

4 Discussion

The present study aimed to investigate the recognition of emotions in children with ASC through the vision of real faces and avatars. The rationale of the research was to understand why avatar-based rehabilitation interventions seem more promising than interventions based on human interaction. Our interest was to understand how ASC children process information about avatars' faces compared to human faces.

Several studies argue that difficulties in the use and comprehension of the information conveyed by human faces could represent a core deficit in children with
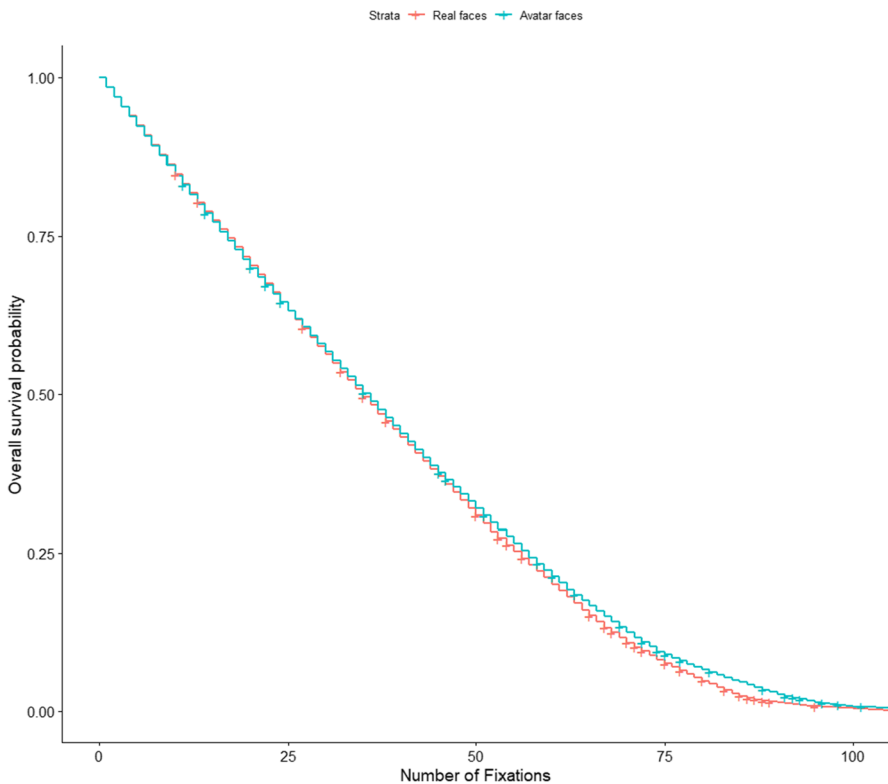

Fig. 4 Survival analysis curves

ASC (Baron-Cohen et al., 1993; Dawson et al., 2002, 2004; Campanelli et al., 2013; Reisinger et al., 2020). But this statement is not entirely true. For example, Castelli (2005) found that children with autism recognized all six primary emotions in the same way as a control sample. Studies of facial emotion recognition in subjects with ASC have yielded mixed results (Harms et al., 2010); thus, the question of whether ASC individuals have general emotion processing impairments remains open to further investigation.

To the best of our knowledge, this is the first study to explore performances between avatar and real faces in a sample of children with ASC. Literature indicates that many factors are related to the recognition of emotions or to visual behaviour when individuals see real faces expressing emotions, these factors include AOIs fixated (mouth or eyes), the emotion expressed by the stimulus (Wegrzyn et al., 2017), the processing style of the individual (Tsang, 2018), and stimulus gender (Lambrecht et al., 2014). Thus, we decided to explore the influences of possible characteristics related to the stimulus.

In our study, we found that children with ASC have less difficulty recognizing the emotion presented by the avatar face as our results showed that the accuracy in emotion recognition was higher with avatar faces than with real ones. This finding supports the use of avatar-based interventions to enhance the abilities of individuals with autism, as they could facilitate emotion recognition.

We gained important insights from the analysis of fixation times; in fact, we found that, compared to real faces, our sample observed with longer fixations the mouth area of avatar faces when the emotion was anger and the eye area when the emotion was sadness, suggesting that these AOIs were of particular interest regarding these emotions, probably enabling enhanced recognition (Fig. 2).

This result is in line with a previous study (Wegrzyn et al., 2017) that found that certain face areas provide greater support in emotion recognition, and it seems that this can also be true for avatar faces. Moreover, we found that fixations were longer on real faces when the emotion was surprise for both the AOIs compared to avatar fixations. This aspect also emerged during the comparison between stimulus type TFDs regarding types of emotions. Specifically, we found that children with ASC fixated for longer on real faces if the emotion was surprise and on avatar faces if the emotions were anger and sadness (Fig. 3), thus it seems that the avatar faces would favour children's attention to negative emotions, and real faces when the emotion is surprise. According to our results, children with ASC spent more time looking at negative emotions in line with evidence that suggests that individuals perform extended scanning for threatening expressions (Green et al., 2003), which has also been found in children with typical development and autism (de Wit et al., 2008). Our results confirm these studies; moreover, it seems that in avatar faces, probably due to better recognition, this behaviour is enhanced compared to real faces. In regard to the emotion of surprise, it has been suggested that recognition of surprise is related to Theory of Mind (Baron-Cohen et al., 1993); in fact, it requires the assessment of another person's mental state (he/she is surprised as he/she was expecting something different), thus it is possible that a simplified stimulus like the one we have constructed may not provide the necessary information for such a complex process. However, further investigation is needed on this aspect. Our results

are particularly intriguing as they provide us with information about the potential use of these stimuli for guiding children's attention and enhancing facial recognition as a potential use, as suggested by Bekele et al. (2014). Since it has been suggested that emotion recognition in ASC children is delayed compared to their peers (Pino et al., 2017) rather than completely absent, a simpler stimulus like the avatar seems to facilitate emotion recognition.

Differences between the two stimuli are particularly important since a combined use of face-to-face and technological interventions has been suggested (Soares et al., 2021), thus it is worth investigating how the two types of stimuli can influence visual processes and promote attention to particular aspects of the social stimulus since social information encoding seems to be of particular relevance in the social cognition of children with ASC (Pino et al., 2020).

Taken together, these results provide an important insight that is useful for clinical practice; in fact, they suggest that the attention of children with ASC could be modulated depending on the stimulus used for a particular emotion, but also concerning a particular area. For example, the research of Dawson and collaborators (Dawson et al., 2005) suggested that people with ASC use atypical strategies for processing faces characterized by reduced attention to the eyes, thus we found that this effect could be reduced depending on the stimulus used. This aspect may be relevant depending on the objectives of the intervention, and it would suggest that the use of both types of stimulus for the purposes of a tailor-made intervention plan may be the best strategy to follow.

In terms of exploration, evaluated by the number of fixations, we found that avatar faces are widely explored by ASC children as they showed a higher number of fixations than real faces, suggesting that avatar faces are more explored than real faces. Even if the causes must be carefully examined, this may be due to the children's interest in the stimulus. In support of this hypothesis, it is worth mentioning the trend towards significance we found during the survival analysis (Fig. 4), where we found that as the number of fixations increased, a greater proportion of children continued to explore avatar faces than real ones. Taken together, these results suggest that avatar faces are more explored than real faces, and in particular, the effects on exploration could be greater in the face of prolonged tasks. This is an important aspect that should be further evaluated, because if we consider the rehabilitation framework, children are encouraged to concentrate on target stimuli several times over a determined period of time (e.g. two/three times a week). Our results show that this topic needs to be more deeply understood as a future perspective.

Despite the interesting results, our study has some limitations as our sample is comprised only of children with ASC; a comparison with a sample of typically developing children could have provided further and important considerations. We decided to use four basic emotions to facilitate their recognition; it is possible that when faced with more complex emotions or scenes, different and equally important results may arise for the intervention framework. As we said before, it is possible that the duration of exposure to the stimulus is another important variable that further studies should examine. In addition, future studies should therefore compare our results with a control group and a more complex paradigm; for example, it would be interesting to use eye-tracking measurements in a real intervention context

to understand in depth a child's behaviour in relation to physical and virtual stimuli. We also believe that considering what interactions there might be when considering more complex emotions could suggest further insights.

In conclusion, we found that a mixture of avatar faces and real faces could be used in order to modulate the attention of a child with ASC on particular facial areas or emotions. Specifically, it seems that avatar faces gain more attention when the emotion is anger or sadness, while real faces gain more attention when the emotion is surprise. AOI attention gains depend on the stimulus type and the emotion expressed by the stimulus. Moreover, we found that avatar faces seem to be more explored than real faces.

Data availability The datasets generated and/or analysed during the current study are not publicly available but they are available from the corresponding author on reasonable request.

Declarations

Ethics approval Ethical approval was obtained from the Ethics Committee of the Reference Regional Centre for Autism. The Ethics Committee approved the experimental protocol number 186061/17 before the recruitment of the participants.

Consent to participate/consent for publication Informed consent was obtained from all parents and oral assent from children before the study.

Conflicts of interest/competing interests The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Amaral, C., Mouga, S., Simões, M., Pereira, H. C., Bernardino, I., Quental, H., et al. (2018). A feasibility clinical trial to improve social attention in autistic Spectrum disorder (ASD) using a brain computer Interface. *Frontiers in Neuroscience, 12*, 477. https://doi.org/10.3389/fnins.2018.00477.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-V* (5th ed.). American Psychiatric Publishing.

Aresti-Bartolome, N., & Garcia-Zapirain, B. (2015). Cognitive rehabilitation system for children with autism spectrum disorder using serious games: A pilot study. *Biomedical Materials and Engineers, 26*(Suppl 1), S811–S824. https://doi.org/10.3233/BME-151373.

Azevedo, J., Silva, V., Soares, F., Pereira, A. P., & Esteves, J. S. (2018). An application to promote emotional skills in children with autism spectrum disorders. In *Joint International Conference on Serious Games* (pp. 282–287). Springer. https://doi.org/10.1007/978-3-030-02762-9_30.

Baron-Cohen, S., Spitz, A., & Cross, P. (1993). Do children with autism recognise surprise? A research note. *Cognition & Emotion, 7*, 507–516. https://doi.org/10.1080/02699939308409202.

Bekele, E., Crittendon, J., Zheng, Z., Swanson, A., Weitlauf, A., Warren, Z., & Sarkar, N. (2014). Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. *Journal of Autism and Developmental Disorders, 44*(7), 1641–1650. https://doi.org/10.1007/s10803-014-2035-8.

Bishop, D. V. M. (2003). *The test for reception of grammar, version 2 (TROG-2)*. Psychological Corporation.

Carter, E. J., Williams, D. L., Hodgins, J. K., & Lehman, J. F. (2014). Are children with autism more responsive to animated characters? A study of interactions with humans and human-controlled avatars. *Journal of Autism and Developmental Disorders, 44*(10), 2475–2485. https://doi.org/10.1007/s10803-014-2116-8.

Caruana, N., Stieglitz Ham, H., Brock, J., Woolgar, A., Kloth, N., Palermo, R., & McArthur, G. (2018). Joint attention difficulties in autistic adults: An interactive eye-tracking study. *Autism, 22*(4), 502–512. https://doi.org/10.1177/1362361316676204.

Cassidy, S. A., Stenger, B., Van Dongen, L., Yanagisawa, K., Anderson, R., Wan, V., et al. (2016). Expressive visual text-to-speech as an assistive technology for individuals with autism spectrum conditions. *Computer Vision and Image Understanding, 148*, 193–200.

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and Normal development. *Autism, 9*(4), 428–449. https://doi.org/10.1177/1362361305056082.

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development, 73*(3), 700–717. https://doi.org/10.1111/1467-8624.00433.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*(2), 271. https://doi.org/10.1037/0012-1649.40.2.271.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology, 27*, 403–424. https://doi.org/10.1207/s15326942dn2703_6.

de Wit, T. C., Falck-Ytter, T., & von Hofsten, C. (2008). Young children with autism spectrum disorder look differently at positive versus negative emotional faces. *Research in Autism Spectrum Disorders, 2*(4), 651–659. https://doi.org/10.1016/j.rasd.2008.01.004.

Di Mascio T., Tarantino L., De Gasperis G., & Pino M.C. (2020). Immersive Virtual Environments: A Comparison of Mixed Reality and Virtual Reality Headsets for ASD Treatment. In: Gennari R. et al. (eds) Methodologies and Intelligent Systems for Technology Enhanced Learning, 9th International Conference. MIS4TEL 2019. *Advances in Intelligent Systems and Computing*, vol 1007. Springer.

Didehbani, N., Allen, T., Kandalaft, M., Krawczyk, D., & Chapman, S. (2016). Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior, 62*, 703–711. https://doi.org/10.1016/j.chb.2016.04.033.

Elgarf, M., Abdennadher, S., & Elshahawy, M. (2017). I-interact: A virtual reality serious game for eye contact improvement for children with social impairment. *Joint International Conference on Serious Games, 10622*, 146–157.

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., Johnson, M. H., et al. (2014). What you see is what you get: Contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience., 9*, 538–543. https://doi.org/10.1093/scan/nst012.

Forbes, P. A., Pan, X., & de C. Hamilton, A. F. (2016). Reduced mimicry to virtual reality avatars in autism Spectrum disorder. *Journal of Autism and Developmental Disorders, 46*(12), 3788–3797.

Franco, F., Itakura, S., Pomorska, K., Abramowski, A., Nikaido, K., & Dimitriou, D. (2014). Can children with autism read emotions from the eyes? The eyes test revisited. *Research in Developmental Disabilities, 35*(5), 1015–1026. https://doi.org/10.1016/j.ridd.2014.01.037.

Green, M. J., Williams, L. M., & Davidson, D. (2003). In the face of danger: Specific viewing strategies for facial expressions of threat? *Cognition and Emotion, 17*, 779–786. https://doi.org/10.1080/02699930302282.

Grossard, C., Chaby, L., Hun, S., Pellerin, H., Bourgeois, J., Dapogny, A., Ding, H., et al. (2018). Children facial expression production: Influence of age, gender, emotion subtype, elicitation condition and culture. *Frontiers in Psychology, 9*, 446. https://doi.org/10.3389/fpsyg.2018.00446.

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism Spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review, 20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6.

Herrero, J. F., & Lorenzo, G. (2020). An immersive virtual reality educational intervention on people with autism spectrum disorders (ASD) for the development of communication skills and problem solving. *Education and Information Technologies, 25*, 1689–1722.

Hopkins, I. M., Gower, M. W., Perez, T. A., Smith, D. S., Amthor, F. R., Wimsatt, F. C., & Biasini, F. J. (2011). Avatar assistant: Improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders, 41*(11), 1543–1555. https://doi.org/10.1007/s10803-011-1179-z.

Jarrold, W., Mundy, P., Gwaltney, M., Bailenson, J., Hatt, N., McIntyre, N., Kim, K., et al. (2013). Social attention in a virtual public speaking task in higher functioning children with autism. *Autism Research, 5*, 393–410. https://doi.org/10.1002/aur.1302.

Jones, E. J. H., Venema, K., Earl, R., Lowy, R., Barnes, K., Estes, A., et al. (2016). Reduced engagement with social stimuli in 6-month-old infants with later autism spectrum disorder: A longitudinal prospective study of infants at high familial risk. *Journal of Neurodevelopmental Disorder., 8*, 7. https://doi.org/10.1186/s11689-016-9139-8.

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders, 43*(1), 34–44. https://doi.org/10.1007/s10803-012-1544-6.

Khowaja, K., Salim, S. S., Asemi, A., Ghulamani, S., & Shah, A. (2019). A systematic review of modalities in computer-based interventions (CBIs) for language comprehension and decoding skills of children with autism spectrum disorder (ASD). *Universal Access in the Information Society*. https://doi.org/10.1007/s10209-019-00646-1.

Kumazaki, H., Warren, Z., Swanson, A., Yoshikawa, Y., Matsumoto, Y., Yoshimura, Y., et al. (2019). Brief report: Evaluating the utility of varied technological agents to elicit social attention from children with autism Spectrum disorders. *Journal of Autism and Developmental Disorders, 49*(4), 1700–1708. https://doi.org/10.1007/s10803-018-3841-1.

Lahiri, U., Bekele, E., Dohrmann, E., Warren, Z., & Sarkar, N. (2015). A physiologically informed virtual reality based social communication system for individuals with autism. *Journal of Autism and Developmental Disorders, 45*(4), 919–931. https://doi.org/10.1007/s10803-014-2240-5.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition and Emotion, 28*(3), 452–469. https://doi.org/10.1080/02699931.2013.837378.

Liu, X. Y., Wu, Q., Zhao, W. B., & Luo, X. (2017). Technology-facilitated diagnosis and treatment of individuals with autism Spectrum disorder: An engineering perspective. *Journal Applied Sciences, 7*(10), 1051. https://doi.org/10.3390/app7101051.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule, 2nd edition (ADOS-2) manual (part I): Modules 1–4*. Western Psychological Services.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces (KDEF)*. CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 91(630), 2–2.

Newbutt, N., Sung, C., Kuo, H. J., Leahy, M. J., & Lin, C. C. (2016). Tong B2Brief report: A pilot study of the use of a virtual reality headset in autism populations. *Journal of Autism and Developmental Disorders, 46*(9), 3166–3176. https://doi.org/10.1007/s10803-016-2830-5.

Orvalho, V., Miranda, J., & Sousa, A. A. (2009). Facial Synthesys of 3D avatars for therapeutic applications. *Studies in Health Technology and Informatics, 144*, 96–98.

Papathomas, P., & Goldschmidt, K. (2017). Utilizing virtual reality and immersion video technology as a focused learning tool for children with autism spectrum disorder. *Journal of Pediatric Nursing, 35*, 8–9. https://doi.org/10.1016/j.pedn.2017.01.013.

Pino, M. C., Mazza, M., Mariano, M., Peretti, S., Dimitriou, D., Masedu, F., et al. (2017). Simple mindreading abilities predict complex theory of mind: Developmental delay in autism Spectrum disorders. *Journal of Autism and Developmental Disorders, 47*(9), 2743–2756. https://doi.org/10.1007/s10803-017-3194-1.

Pino, M. C., Vagnetti, R., Masedu, F., Attanasio, M., Tiberti, S., Valenti, M., & Mazza, M. (2020). Mapping the Network of Social Cognition Domains in Children With Autism Spectrum Disorder Through Graph Analysis. *Frontiers in Psychiatry*, 11. https://doi.org/10.3389/fpsyt.2020.579339.

Ravindran, V., Osgood, M., Sazawal, V., Solorzano, R., & Turnacioglu, S. (2019). Virtual reality support for joint attention using the floreo joint attentionmodule: Usability and feasibility pilot study. *JMIR Pediatrics and Parenting, 2*(2), e14429. https://doi.org/10.2196/14429.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reisinger, D. L., Shaffer, R. C., Horn, P. S., Hong, M. P., Pedapati, E. V., Dominick, K. C., & Erickson, C. A. (2020). Atypical social attention and emotional face processing in autism Spectrum disorder: Insights from face scanning and Pupillometry. *Frontiers in Integrative Neuroscience, 13*, 76. https://doi.org/10.3389/fnint.2019.00076.

Rojo, D., Mayor, J., Rueda, J. J. G., & Raya, L. (2019). A virtual reality training application for adults with Asperger's syndrome. *Journal - IEEE Computer Graphics and Applications, 39*(2), 104–111. https://doi.org/10.1109/MCG.2018.2884272.

Santos, P., Silva, V., Soares, F., & Simões, A. (2019). Facial virtual tracking: a system to mirror emotions. In *EPIA Conference on Artificial Intelligence* (pp. 68–79). Springer. https://doi.org/10.1007/978-3-030-30244-3_7.

Singular Inversions Inc. (2009). FaceGen Modeller (Version 3.2) [Computer Software]. retrieved from http://www.facegen.com.

Shoaib, M., Hussain, I., Mirza, H. T., & Tayyab, M. (2017). The role of information and innovative technology for rehabilitation of children with Autism: A systematic literature review. *In 2017 17th International Conference on Computational Science and Its Applications (ICCSA),* (pp. 1–10). IEEE. https://doi.org/10.1109/ICCSA.2017.7999647.

Soares, E. E., Bausback, K., Beard, C. L., Higinbotham, M., Bunge, E. L., & Gengoux, G. W. (2021). Social skills training for autism Spectrum disorder: A meta-analysis of in-person and technological interventions. *Journal of Technology in Behavioral Science, 6*(1), 166–180. https://doi.org/10.1007/s41347-020-00177-0.

Suraniti, S., Ferri, R., & Neri, V. (2009). Test for reception of grammar: TROG-2. Curatori edizione italiana. Giunti O. S.

Tanaka, H., Negoro, H., Iwasaka, H., & Nakamura, S. (2017). Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PLoS One, 12*(8), e0182151. https://doi.org/10.1371/journal.pone.0182151 eCollection 2017.

Therneau T (2020). *A Package for Survival Analysis in R*. R package version 3.2–7, https://CRAN.R-project.org/package=survival.

Tsang, V. (2018). Eye-tracking study on facial emotion recognition tasks in individuals with high-functioning autism spectrum disorders. *Autism, 22*(2), 161–170. https://doi.org/10.1177/1362361316667830.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5.

Valencia, K., Rusu, C., Quiñones, D., & Jamet, E. (2019). The Impact of Technology on People with Autism Spectrum Disorder: A Systematic Literature Review. *Journal Sensors*, *19*(20). https://doi.org/10.3390/s19204485.

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS One, 12*(5), e0177239.

Yuan, S. N. V., & Ip, H. H. S. (2018). Using virtual reality to train emotional and social skills in children with autism spectrum disorder. *London Journal of Primary Care, 10*(4), 110–112. https://doi.org/10.1080/17571472.2018.1483000.

Authors and Affiliations

Maria Chiara Pino[1,2] · Roberto Vagnetti[1] · Marco Valenti[1,2] · Monica Mazza[1,2]

> Maria Chiara Pino
> mariachiara.pino@univaq.it

> Marco Valenti
> marco.valenti@univaq.it

> Monica Mazza
> monica.mazza@univaq.it

[1]   Department of Applied Clinical Sciences and Biotechnology, University of L'Aquila, Via Vetoio, Località Coppito, 67100 L'Aquila, Italy

[2]   Regional Reference Centre for Autism of the Abruzzo Region, Local Health Unit ASL 1, L'Aquila, Italy