Check for updates

Received: 26 May 2021 | Accepted: 28 October 2021

DOI: 10.1002/aur.2636

RESEARCH ARTICLE

Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis

Mariano Alcañiz[1] | **Irene Alice Chicchi-Giglioli**[1] | **Lucía A. Carrasco-Ribelles**[1,2] | **Javier Marín-Morales**[1] | **Maria Eleonora Minissi**[1] | **Gonzalo Teruel-García**[1] | **Marian Sirera**[3] | **Luis Abad**[3]

[1]Instituto de Investigación e Innovación en Bioingeniería (i3B), Universitat Politécnica de Valencia, Valencia, Spain

[2]Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain

[3]Red Cenit, Centros de Desarrollo Cognitivo, Valencia, Spain

Correspondence
Irene Alice Chicchi-Giglioli, Instituto de Investigación e Innovación en Bioingeniería (i3B), Universitat Politécnica de Valencia, Valencia, Spain.
Email: alicechicchi@i3b.upv.es

Funding information
Spanish Ministry of Economy, Industry, and Competitiveness-funded project "Immersive Virtual Environment for the Evaluation and Training of Children with Autism Spectrum Disorder: T Room, Grant/Award Number: IDI-20170912

Abstract

The core symptoms of autism spectrum disorder (ASD) mainly relate to social communication and interactions. ASD assessment involves expert observations in neutral settings, which introduces limitations and biases related to lack of objectivity and does not capture performance in real-world settings. To overcome these limitations, advances in technologies (e.g., virtual reality) and sensors (e.g., eye-tracking tools) have been used to create realistic simulated environments and track eye movements, enriching assessments with more objective data than can be obtained via traditional measures. This study aimed to distinguish between autistic and typically developing children using visual attention behaviors through an eye-tracking paradigm in a virtual environment as a measure of attunement to and extraction of socially relevant information. The 55 children participated. Autistic children presented a higher number of frames, both overall and per scenario, and showed higher visual preferences for adults over children, as well as specific preferences for adults' rather than children's faces on which looked more at bodies. A set of multivariate supervised machine learning models were developed using recursive feature selection to recognize ASD based on extracted eye gaze features. The models achieved up to 86% accuracy (sensitivity = 91%) in recognizing autistic children. Our results should be taken as preliminary due to the relatively small sample size and the lack of an external replication dataset. However, to our knowledge, this constitutes a first proof of concept in the combined use of virtual reality, eye-tracking tools, and machine learning for ASD recognition.

Lay Summary

Core symptoms in children with ASD involve social communication and interaction. ASD assessment includes expert observations in neutral settings, which show limitations and biases related to lack of objectivity and do not capture performance in real settings. To overcome these limitations, this work aimed to distinguish between autistic and typically developing children in visual attention behaviors through an eye-tracking paradigm in a virtual environment as a measure of attunement to, and extraction of, socially relevant information.

KEYWORDS
autism spectrum disorder, behavioral biomarker, eye tracking, machine learning, multivariate supervised learning, virtual reality

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Autism Research* published by International Society for Autism Research and Wiley Periodicals LLC.

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with an estimated worldwide prevalence of 1 in 160 among children (World Health Organization [WHO, 2019]). The DSM-V and ICD-11 are the two international gold-standard classification manuals that provide criteria for diagnosing ASD (American Psychiatric Association, 2013; WHO, 2019). According to the DSM-V and ICD-11, the main symptoms of ASD concern impairments in social and interaction abilities and the presence of restrictive interests and repetitive behaviors. Symptom onset typically occurs between the ages of 2 and 4 years, although in some cases the first symptoms can occur as early as 6 months (mainly related to abnormalities in eye contact and language development, followed by failure to initiate or respond to social interactions and difficulty understanding others' intentions in social contexts). There are two standardized tools for ASD assessment: semi-structured observational tasks for children (the Autism Diagnostic Observation Schedule [ADOS-2]; Lord et al., 1999) and a semi-structured interview for parents (the Autism Diagnostic Interview-Revised [ADI-R]; Lord et al., 1994). Evaluation relies on children's observable behaviors and parents' interview responses, which clinicians rate according to their expertise and subjectivity. Although these measures are well validated, the qualitative methodologies have some limitations and biases that can provide inaccurate and/or misleading outcomes and interpretations. In fact, the main limitation for researchers and clinicians concerns the lack of objective methods for assessment, since the actual evaluation includes qualitative clinical observations of manifest symptoms, mainly related to social, communicative, and interactive abilities (Lord et al., 1999, 2001). Furthermore, assessment occurs in laboratories or clinical settings lacking ecological validity that could offer a deeper understanding of real-life abilities. Finally, social desirability bias has been found to affect the veracity of parents' responses, according to a favorable view by others.

To overcome these limitations, clinical research has attempted to identify more quantifiable and objective characteristics of biological and unconscious ASD processes, also known as biomarkers. Objectifying and quantifying unconscious processes could provide a more systemic diagnosis and earlier detection of ASD, allowing for more customized and earlier interventions. Recent advances in implicit measures and tools (i.e., electrodermal activity, body movements, eye tracking) and technological systems (i.e., virtual reality [VR]) have enabled the capture of unconscious processes for the identification of biomarkers from a dimensional perspective, enriching assessments with more objective data than conventional assessments typically contain, and the creation of ecologically valid environments that can provide dynamic stimuli that resemble real life, gathering performance in real time (Alcañiz, Chicchi Giglioli, et al., 2020; Alcañiz, Marín-Morales, et al., 2020).

Implicit approaches and measures in ASD assessment

Biomarkers refer to unconscious indicators that can potentially be used to identify multiple unobservable processes in ASD, especially in highly heterogeneous conditions, enabling the improvement of diagnosis and recognition of subgroups. Genetic, neural, physiological, and behavioral characteristics are the main biomarkers that have been identified and investigated in ASD (Bridgemohan et al., 2019; Del et al., 2018; Ruggeri et al., 2014). Regarding brain activation and neural activity, functional magnetic resonance imaging studies on social contexts have differentiated those with ASD from the typically developing (TD) population in terms of prefrontal cortex activity, and electroencephalography studies have shown different neural responses to social and nonsocial stimuli over the occipital cortex for the two populations (Sumiya et al., 2020; Vettori, Dzhelyova, Van der Donck, Jacques, Van Wesemael, et al., 2020a; Vettori, Dzhelyova, Van der Donck, Jacques, Steyaert, et al., 2020b).

Regarding studies on physiological biomarkers (including electrodermal activity and heart rate variability) in social contexts, recent studies have demonstrated an accuracy of 85% in differentiating between autistic and TD children (Alcañiz, Chicchi Giglioli, et al., 2020). Two relevant biomarkers for behavioral responses in social contexts are body movement recognition obtained using accelerometers or cameras with depth sensors (i.e., RGB-D) and gaze behavior obtained using eye-tracking tools (i.e., Tobii Pro Glasses 2; Falck-Ytter et al., 2015; Gonçalves et al., 2012; Min & Tewfik, 2010; Thorup et al., 2016). Body movement recognition studies have effectively identified repetitive behaviors in autistic children—mainly related to the head, trunk, and feet—with an accuracy of 82.98% (Alcañiz, Marín-Morales, et al., 2020). On the other, eye gaze behavior has proven and continues to be the most relevant biomarker for autistic children due to its feasibility and non-intrusiveness, which allow child development abnormalities to be detected earlier than is possible with conventional recognition tests (Falck-Ytter et al., 2015; Thorup et al., 2016).

Eye gaze to social attentional cue recognition in ASD

Social situations require various abilities related to social information processing, such as face emotional recognition, social play, exchanges, and comprehension of others' intentions and goals. These abilities mainly

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

depend on social attentional cue abilities, including joint attention; paralinguistic cues, such as body posture and movement, head orientation, and hand gestures; and linguistic verbalizations. Studies of children that have relied on conventional observational tasks (i.e., the ADOS) have shown that children with ASD showed atypical patterns of attention to social cues, characterized by less attention to faces, people, and social situations compared with TD children. Eye-tracking systems allow researchers to detect eye movement and analyze areas of interest (AOIs), enabling more objective assessment of social attentional cue abilities compared with traditional methods. Chita-Tegmark's (2016a) meta-analysis of eye-tracking studies on six AOIs (eyes, mouth, face, body, nonsocial stimuli, and background) comparing behaviors between autistic and TD children showed that autistic children spent less time than TD children looking at eyes, mouths, and faces in social stimuli conditions and more time looking at bodies. Furthermore, autistic children spent more time looking at nonsocial stimuli, such as the background, than social stimuli, but no differences were found in comparison with TD. Several studies, however, did not find the eye gaze behavior in ASD that Chita-Tegmark (2016b) reported. For instance, autistic children looked at the eyes for the same amount of time as TD children in static social stimuli (e.g., de Wit et al., 2008; Rutherford & Towns, 2008; van der Geest et al., 2002) and looked equally at social and nonsocial elements in dynamic stimuli (e.g., Parish-Morris et al., 2013). Such differences might be dependent on the study methodology and the type of social stimuli involved. Using either static or dynamic social stimuli might yield different results, and only a few studies have compared these two conditions (e.g., Chevallier et al., 2015; Cilia et al., 2019; Saitovich et al., 2013; Shic et al., 2014; Speer et al., 2007). Shic et al. (2014) compared a neutral female face image, a video with a woman smiling, and a video with a woman smiling and speaking and found that autistic children looked less often at the eyes in the social stimuli condition and in response to the dynamic stimulus of the woman speaking. Chevallier et al. (2015) compared static and dynamic visual tasks involving both social and nonsocial stimuli and found that autistic children spent less time looking at social than nonsocial stimuli compared with TD children. Finally, a recent study by Cilia et al. (2019) showed that both static and dynamic stimuli were relevant in distinguishing autistic from TD children using eye gaze. On the one hand, static stimuli enabled AOIs to be identified with greater precision, showing similar patterns in the two populations. On the other, dynamic stimuli could better discriminate among various modalities of social interaction (i.e., pointing, head orientation, verbalization) in autistic children, highlighting that pointing is the most relevant element in guiding children's visual attention.

In addition, particular differences between groups in eye gaze behavior in response to static versus dynamic social stimuli, a new trend of research is emerging with regard to the objective assessment of ASD based on social visual attention and machine learning (ML) techniques (Minissi et al., 2021). Liu et al.'s (2016) pioneering study identified autistic children based on eye gaze in response to static social stimuli with an accuracy of 88.51%, sensitivity of 93.10%, specificity of 86.21%, and AUC of 0.89. Likewise, He et al. (2021) achieved 81.1% accuracy in the classification of TD children, low-functioning autistic children, and high-functioning autistic children based on eye movements during a visual-orienting task involving static stimuli with gaze-related or non–gaze-related directional cues. In addition to static stimuli, the combination of eye movements in response to dynamic social stimuli with ML techniques has proven effective in the early discrimination and classification of ASD (e.g., Carette et al., 2017, 2019; Wan et al., 2019). The aforementioned studies, as well as the majority of studies to date on the eye gaze behavior of autistic children in response to social stimuli, involved so-called "offline" social cognition: the use of static or dynamic social stimuli presented on desktop devices, lacking direct social interaction that resembles real-world contexts (Schilbach, 2014). Whether autistic children would present the same eye gaze behavior in real or realistic social contexts is still unclear, since further evidence has postulated that findings related to offline social cognition cannot be generalized across contexts (Guillon et al., 2014). Based on those findings, researchers began to use head-mounted eye-tracking systems to clarify the eye gaze behavior of autistic children in real social interaction contexts as a measure of "online" social cognition (Schaller et al., 2021; Zhao et al., 2021). Head-mounted eye-tracking systems allow eye gaze behaviors to be measured in both real and realistic situations; they have demonstrated suitability for autistic children (e.g., Schaller et al., 2021; Zhao et al., 2021) and hence feasibility for assisting in early detection of ASD, enhancing its ecological validity. However, real settings might be challenging for autistic children, due to their sensory dysfunction and impairments in daily life skills. Realistic situations, such as those provided by VR systems, can ensure both an ecologically valid setting and a controlled environment wherein it is safe to either test or train ASD participants.

Based on the previous literature involving eye-tracking studies and the above-mentioned methodological assessment limitations in ASD, recent advances in technologies like VR can deliver realistic simulated situations characterized by a high sense of presence and ecological validity in which static and dynamic stimuli can be tuned and controlled while participants' eye gaze behavior is recorded.

Virtual reality in ASD assessment

VR can be defined as a three-dimensional synthetic system in which realistic simulated environments can be

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

developed (Burdea, 2003). VR systems provide immersion—the technological capacity to isolate the user from reality, which depends on the device and human–computer interaction—through control sticks or gloves with which the user can interact with the virtual objects. VR also provides sense of presence: the psychological feeling of "being in" the virtual scenario, as if the user were in the real world (Cipresso et al., 2018; Slater et al., 2009). Thanks to these features, VR systems offer ecologically valid environments characterized by engagement, motivation, fun, and the ability to gather behavioral performance during gameplay. In ASD, various VR applications have been tested related to both treatment and assessment. Regarding ASD treatment, training programs on desktop devices for social competences, emotional recognition, anxiety, and phobias using implicit measures, such as eye tracking, have shown effectiveness and improvement in ASD populations (Parsons, 2016; Parsons & Mitchell, 2002). Regarding the assessment and diagnosis of ASD, immersive VR and implicit biomarkers have been less addressed and are currently starting to prove their effectiveness (Alcañiz, Chicchi Giglioli, et al., 2020; Alcañiz, Marín-Morales, et al., 2020). To our knowledge, no studies have investigated the feasibility of early assessment of ASD based on ML techniques and eye gaze in response to social versus non-social cues presented in an immersive VR environment.

There are promising findings in ASD classification based on ML techniques and offline social cognition (i.e., social visual attention in desktop devices; e.g., Carette et al., 2017, 2019; He et al., 2021; Liu et al., 2016; Wan et al., 2019). These results offer a powerful reason to attempt the application of the same methodology in more controlled, realistic, and ecological settings. Indeed, static, and dynamic social stimuli presented on desktop devices (i.e., offline social cognition), although effective, differ from reality in many aspects, and the involvement of immersive VR might lead to more objective results due to the superior, more realistic user experience it provides (Minissi et al., 2021). Autistic children, however, may not accept immersive VR systems such as head-mounted displays, since they may exacerbate sensory and cognitive difficulties and may not fit on a child's head (Guazzaroni et al., 2018; Wallace et al., 2010). In this context, semi-immersive VR systems represent a feasible solution. For example, the Cave Assisted Virtual Environment (CAVE™), which has already proven feasibility with autistic children (e.g., Alcañiz, Chicchi Giglioli, et al., 2020; Alcañiz, Marín-Morales, et al., 2020; Cai et al., 2013), offers a safe environment in which users can experience and interact with realistic virtual elements without needing to wear VR helmets. Starting from this premise, the primary aim of this study was to distinguish autistic from TD children in visual attention behaviors through an eye-tracking paradigm in a virtual environment as a measure of attunement to, and extraction of, socially relevant information. Specifically, we explored (1) whether it is possible to distinguish between the two populations using eye gaze data and (2) which parameters better distinguish the two populations.

METHODS

Participants

The 55 children aged between 4 and 7 years participated in the study: 20 TD children (M_{age} = 4.75 years, SD = 0.77) and 35 diagnosed with ASD (M_{age} = 5.26 years, SD = 0.51). Autistic children were recruited from the Red Cenit Neurocognitive Development Center, Valencia, Spain. The TD group was recruited by a management company through calls and mailings to families. Both groups were individually evaluated with the same scales and procedure prior to the experiments. To participate in the study, participants were required to not wear glasses or present any alteration or ocular pathology. Prior to inclusion in the study, the relatives of all participants received and signed an informed consent form explaining the objectives of the research and the characteristics of the experimental procedure. They also consented to video recording of the participating subject. The study obtained the approval of the Ethics Committee at the Polytechnic University of Valencia, and the entire procedure was designed following the guidelines of the Declaration of regarding the ethical standards to be followed in any procedure that includes human beings.

Psychological assessment

The evaluation protocol consisted of the following diagnostic tests.

The ADOS-2 (Lord et al., 1999) is a semi-structured scale that includes different tasks. Its objective is to evaluate children's development in various areas, such as social interaction and play, in order to observe possible symptoms of autism, such as communication deficits and/or the presence of restrictive and repetitive behaviors. The ADOS-2 contains five modules designed to evaluate a wide range of the population in terms of age (31 months through adolescence and adulthood) and linguistic level (ranging from absence of phrase language to fluent language). A trained psychologist observes and scores the different behaviors to obtain two specific indexes (social impairment and restricted and repetitive behavior) along with a global total index of ASD. The ADOS-2 scale has high test–retest reliability (0.87 for the social impairment index, 0.64 for the repetitive behavior index, and 0.88 for the total global index), making it the test par excellence of ASD diagnosis. In this study, the evaluation was carried out using module 1, which corresponds to children 31 months of age and older who do not use coherent phrase language.

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The ADI-R (Lord et al., 1994) is a semi-structured interview oriented toward and answered by relatives of children and adults with suspected ASD. Its objective is to provide a framework of history of development from childhood throughout life to detect the presence of ASD symptoms. The 111 questions are scored on a Likert scale ranging from 0 to 3, following the criteria and separation established by the ICD-10 and DSM-IV. Three indices are obtained: communication; social interaction; and restricted, repetitive, and stereotyped behaviors. The test–retest reliability of the ADI-R ranges from 0.93 to 0.97, making it an effective tool with excellent psychometric properties.

Virtual environment

The virtual content was designed by the Institute for Research and Innovation in Bioengineering (i3b) of the Polytechnic University of Valencia. The virtual environment was designed to be projected in a 2D three-wall CAVE™ system without stereoscopy and perspective correction and with dimensions of 4 m × 4 m × 3 m. Its equipment consisted of three ultrashort lens projectors (visual component), Logitech Speaker System Z906599W 5.1 HX digital speakers (auditory component), and a wireless Olorama™ system (https://www.olorama.com) that regulated the presence and intensity of different odors (olfactory component; Figure 1).

The virtual experience took place in a mall composed of various shopping and entertainment stores (central hall, electronics shop, game center, supermarket, and cinema) and stimuli (train, carousel, trash, exhibitors, line boxes, and billboard) in which the participant was stimulated visually, auditorily, and olfactorily (Table 1; Figure 2a, b). The virtual experience was characterized by various static and dynamic social and nonsocial stimuli as well as various virtual agents (children and adults) who interacted with the participant to explore the virtual mall. The duration of the experience was 24 min and 45 s for each subject.

Experimental procedure

First, participants' relatives were informed about the general objectives of the research. Before the experimental session, the researchers showed and explained the environment of the experimentation to them. Regarding the experimental session, eye-tracking glasses were first placed in a room next to the 2D CAVE™, where they were calibrated and the researchers verified that they operated correctly and the subject did not reject them. Subsequently, recording began and the participant was led into the CAVE™, where they were placed in the center of the room, standing 1.5 m away from the central wall (except in the last scene, where they sat on the floor). Although the virtual experience always began in the presence of the researcher in order to monitor the child's behavior, the researchers attempted to intervene as little as possible (only in situations of device failure or cybersickness). To avoid cognitive and sensory overload, the participant was presented with a scene where a forest appeared with relaxing music, both at the beginning of the experiment and in transitions between stores. The order of the presentation of the virtual scenes was counterbalanced among participants.

Eye gaze assessment and data processing

Data on each participant's eye gaze were collected using Tobii Pro Glasses 2 (https://www.tobiipro.com/product-

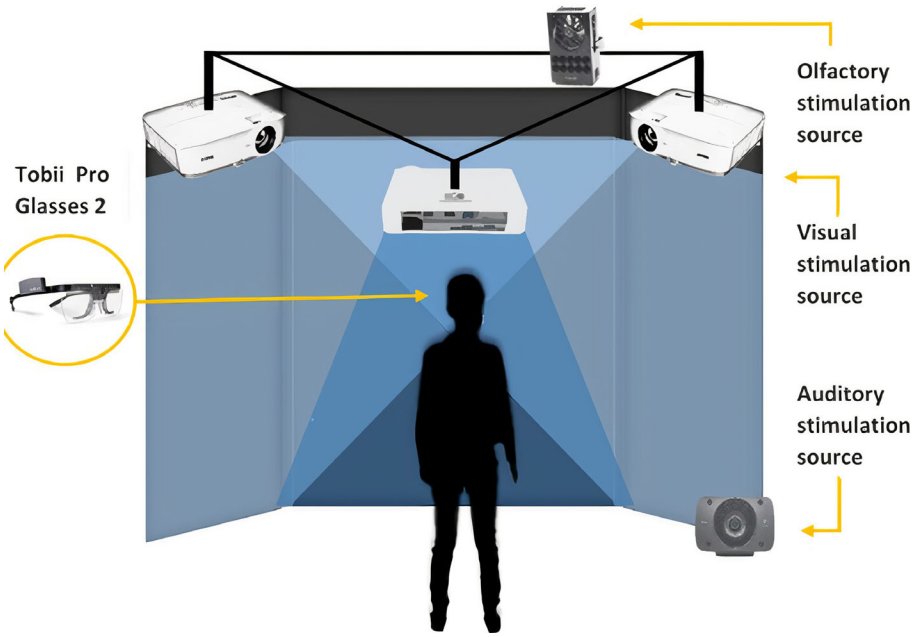

FIGURE 1 Experimental setting

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Description and duration of each virtual scenario

Countdown	A countdown appears to anticipate the start of the virtual experience.	10 s
Central hall	The avatar welcomes the participant, waves at them, briefly explains what they will see, and asks for their name.	180 s
Train	An illuminated train appears and disappears four times in the central hall; the first time, it almost runs over the avatar.	180 s
Carousel	A merry-go-round appears in motion. There are some children riding the merry-go-round, and the avatar is raised. Bubbles appear in the scene twice.	90 s
Trash	A trash can overflowing with trash appears, and the avatar shows signs of disgust.	45 s
Electronics shop	The avatar appears at the door of an electronics store and asks the participant if they want to accompany him to look for computer games. Subsequently, he urges the participant to enter.	30 s
Clerk	A clerk appears in the store hall and disappears off to one side.	35 s
Exhibitors	The avatar hides in the displays and plays with the participant by appearing and disappearing four times.	30 s
Supermarket door	The avatar appears at the door of a supermarket and asks the participant if they want to accompany him in looking for his favorite goodies. Later, he urges the participant to enter.	10 s
Line boxes	The avatar appears with an apple in one hand and a lollipop in the other and asks the participant which of the two they want.	25 s
Billboard	A billboard appears. The ads change, and the avatar urges the participant to look at the billboard. Later, he urges the participant to continue seeing the mall.	20 s
Game room door	The avatar appears at the door of a game room and shows great interest in playing. Later, he urges the participant to enter.	40 s
Game room	The avatar appears and indicates three times, both verbally and with a hand gesture, that the participant should look to the left. A ball then appears and rotates to the right until it disappears.	40 s
Cinema door	The avatar appears, indicates that the cinema is showing his favorite movie that week, and asks the participant if they want to see it with him. Later, he urges the participant to enter.	35 s
Cinema seats	A movie theater appears with two adult avatars on the side screens and the first scene of the movie *Inside Out* projected on the screen.	45 s
Countdown	A countdown appears to anticipate the end of the virtual experience.	10 s

FIGURE 2 (a) Carousel and (b) Exhibitors in the electronics shop

listing/tobii-pro-glasses-2), an eyeglass device that records in first person what the participant is observing while moving freely through space. These recordings are an ideal source of direct and objective information for studying eye gaze behavior. The device is equipped with a microphone, a front camera facing the external environment, and two cameras for each eye that use a 3D eye model that enables eye-tracking studies in dynamic environments. It is also equipped with an accelerometer and a gyroscope that allow for differentiation between head and eye movements, eliminating the impact of head movements on tracking data. The recordings were subsequently treated using an ad hoc program consisting of synchronization among the recorded videos of each participant, the patterns of fixations, the frame-by-frame starting images of the virtual environment, and the data referring to the previously defined AOIs. As a result of this process, a text file reporting the frames in which each AOI was seen was obtained. The frames in which the participant was not looking at any defined AOI were not reported. The scenes defined in Table 1 were categorized into seven groups: game room, cinema, electronics shop, supermarket, forest, central hall, and carousel. The

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

defined AOIs were grouped into categories and subcategories of interest: people (children and adults), faces and bodies of people (children and adults), items (dynamic and static), and background. The features created were related to each point described in Figure 3 and were extracted, both for the full experience and for each scene. Ultimately, 144 features were extracted. Some examples are as follows:

- General features: Number of AOIs seen, average number of AOIs seen per scene (and standard deviation), and so on.
- Background: Number of frames in which the participant did not see anything defined as an AOI.
- AOIs: Number of frames in which the participant did see something defined as an AOI.
- Persons: Number of frames in which the participant saw any defined character, as well as (for example) the number of frames in which the avatar was seen, the number of frames in which the rest of the characters were seen, the number of frames in which the participant saw faces, and the number of frames in which the participant saw bodies.
- Items: Number of frames in which the participant saw any defined item. The number of frames in which the participant saw dynamic items (e.g., carousel) and static items (e.g., trash) were also calculated as independent variables.

All features except general features were calculated for the experience as a whole as well as for each scene. Some extra features were also created to define the difference between some of the variables mentioned above, such as the difference in the number of frames in which faces were looked at compared with bodies (i.e., in how many more frames were faces looked at instead of bodies), the difference between children and adults, the difference between the avatar and other characters, the difference between dynamic and static items, and the difference between people and items.

Data analysis

First, hypothesis testing was performed to find variables that showed statistically significant differences between participants with ASD and TD participants. A *t*-test was used for normally distributed variables and a Mann–Whitney test for non-normally distributed variables. The normality of each variable was tested before using the Shapiro–Wilk test. Variables with a p value lower than 0.05 were considered not normally distributed. Level of statistical significance was set as $\alpha < 0.05$.

A set of ML models were made to further study the influence of each variable on the diagnosis of ASD, creating different datasets. Two approaches were followed: A general and a hypothesis contrast approach.

In the general approach, two datasets were used:

1. Dataset with all available features (i.e., 144).
2. Dataset with only those features in which statistically significant differences were previously found.

In the hypothesis contrast approach, several datasets were used to test the influence of each set of variables:

a. Variables for the number of frames in which people in general and items were seen, in the experience as a whole and per scene (23 features).

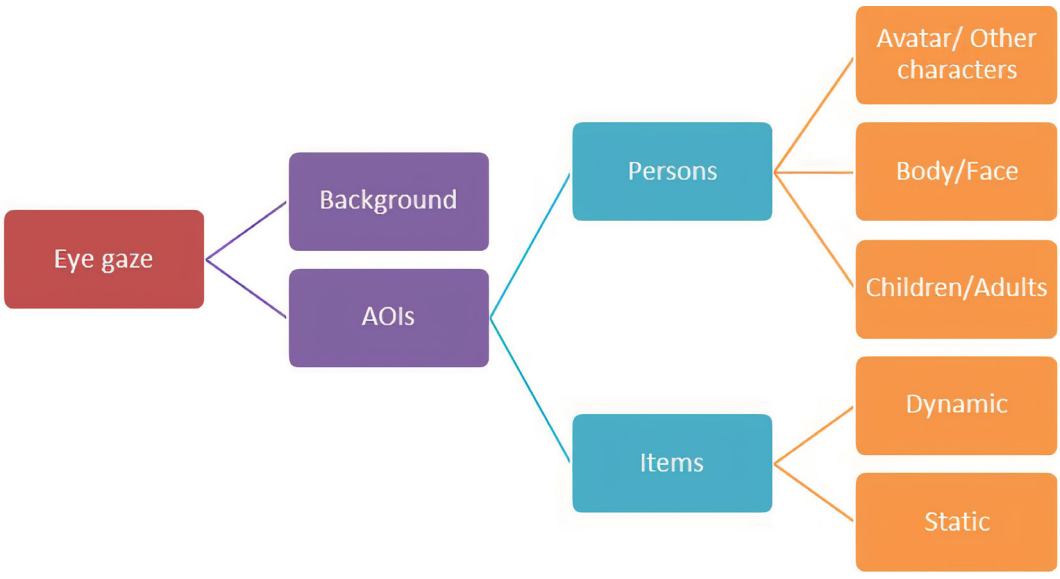

FIGURE 3 Scheme of features created using eye gaze information

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

b. Variables for the number of frames in which children and items were seen, in the experience as a whole and per scene (17 features).
c. Variables for the number of frames in which adults and items were seen, in the experience as a whole and per scene (17 features).
d. Variables for the number of frames in which adults and children were seen, in the experience as a whole and per scene (10 features).
e. Variables for the number of frames in which faces and bodies were seen, in the experience as a whole and per scene (22 features).
f. Variables for the number of frames in which children's faces and bodies were seen, in the experience as a whole and per scene (eight features).
g. Variables for the number of frames in which adult faces and bodies were seen, in the experience as a whole and per scene (seven features).
h. Variables for the number of frames in which dynamic and static items were seen, in the experience as a whole and per scene (16 features).

The number of features in each dataset is not consistent, as not all scenes had all AOIs (e.g., some did not show any adults or any dynamic items).

For each dataset, feature selection was performed using a step backward sequential wrapper. A maximum number of features to be selected was set to build predictive models with up to 10 features in order to avoid overfitting. Both the best feature subset and the best model were chosen using five-fold cross-validation (repeated four times). The average of the following metrics is reported, along with their standard deviations: accuracy, kappa, AUC, F_1 score, sensitivity (TPR), and specificity (TNR).

For each dataset, the following algorithms were trained as previously described: naïve Bayes, XGBoost, kNN, random forest, and SVM.

All the analyses were performed in R (version 3.6.1). ML was performed using the R package mlr (Bischl et al., 2016). The models were trained using a PC with an eight-core Intel Core i7-8700F CPU and 16 GB RAM.

RESULTS

Eye-tracking analysis

A total of 13 variables showed statistically significant differences between autistic and TD participants. Figure 4 shows their distribution. Two of these variables are related to the number of AOIs seen throughout the experience (Figure 4, boxplots 1 and 2). Three of them are related to items in the game center: autistic participants watched significantly more items than TD participants and, in particular, more static items. This is also reflected in the differences in the number of frames in which participants watched dynamic items rather than static items, which was higher for autistic participants. TD participants, on the other hand, focused more on dynamic items (Figure 4, boxplots 3–5). In the hall, autistic participants looked at adults and characters other than the avatar in significantly more frames than TD participants (Figure 4, boxplots 6 and 13). In the electronics shop, autistic participants looked at other children (both their faces and their bodies) in significantly more frames than did TD participants (Figure 4, boxplots 8–10). They also looked at the main avatar in more frames than did TD participants (Figure 4, boxplot 7). Autistic participants looked at characters' faces in significantly more frames than did TD participants in the electronics shop and game center (Figure 4, boxplots 11 and 12).

ASD recognition models

Table 2 shows the results of the general approach ML models. In this approach, two datasets (one with all variables and another with only the 13 variables that had significant results in the statistical analysis) were used to fit models using the pipeline described in the data analysis section. The main result demonstrated 86% accuracy with 91% sensitivity in the recognition of autistic children when using all eye-tracking variables.

Table 3 shows the results of the ML models' approach regarding the specific parameters that hypothetically could better discriminate between the two populations. In this approach, eight different datasets were tested to fit models, as described in the data analysis section. These datasets sought to study the influence of more specific sets of variables, specifically: (a) number of frames of people (adults and children) and static items; (b) number of frames of children and items; (c) number of frames of adults and items; (d) number of frames of children and adults; (e) number of frames of face and body; (f) number of frames of children's faces and bodies; (g) number of frames of adults' faces and bodies; and (h) number of frames of dynamic and static items.

The main results in the recognition of autistic children were found in the (a), (b), (c), and (d) models, which demonstrated accuracy levels of 84% (94% sensitivity), 82% (92% sensitivity), 83% (82% sensitivity), and 78% (90% sensitivity), respectively.

Note 1 AOIs: total number of visited/seen AOIs; 2. NavgAOIperScene: average number of AOIs seen per scene; 3. NFramesItems_GameCenter: total number of stimuli frames seen in the game center; 4. NFramesItems Static_GameCenter: number of static stimuli seen in the game center; 5. NFramesPersons Adults_Hall: number of frames of adults seen in the hall; 6. NFramesPersonsAvatar_ElectronicShop: number of frames of main avatar

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

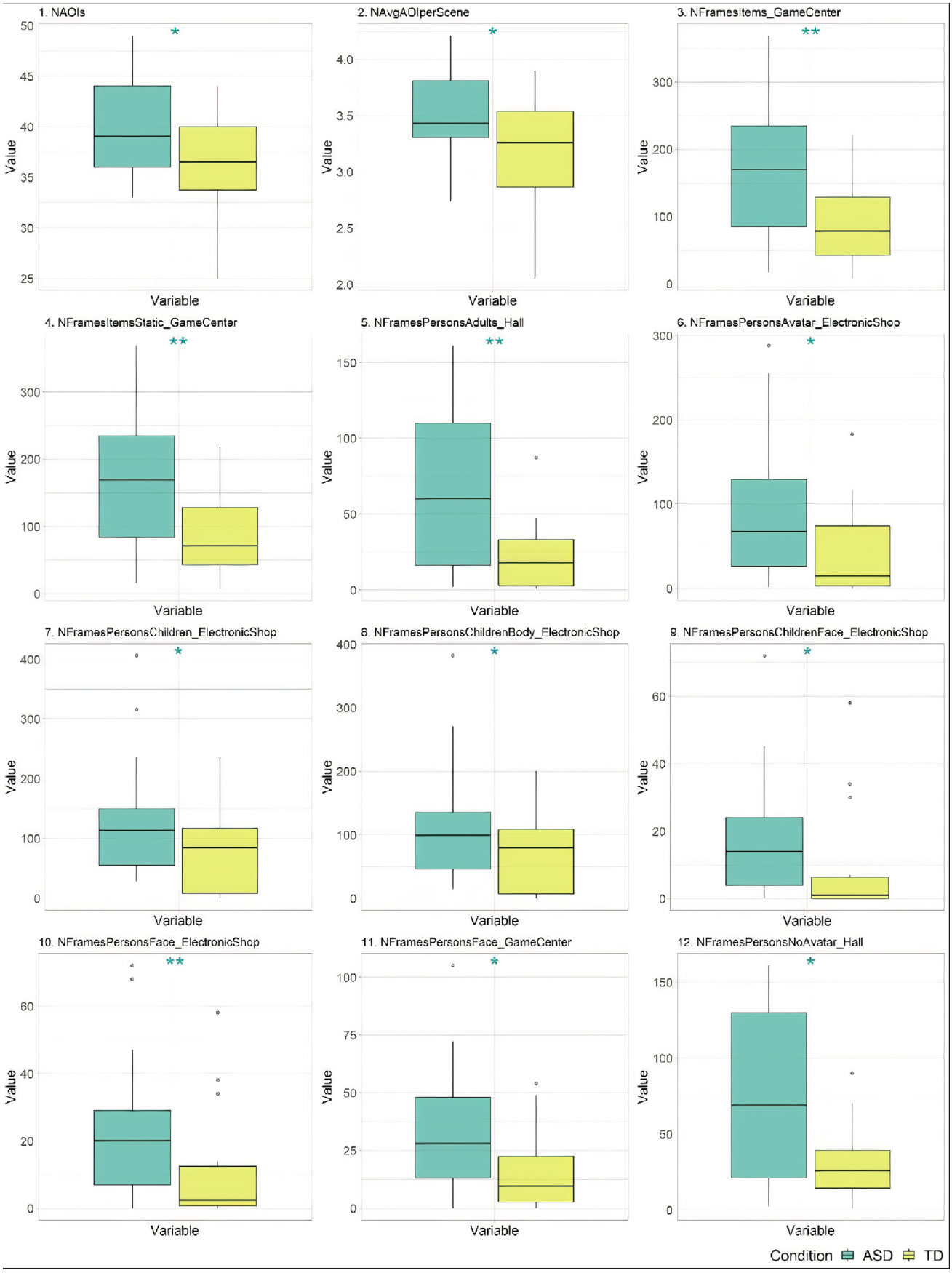

FIGURE 4 Boxplots of statistically significant differences in the distribution of ASD and TD children

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Results of general approach ML models. The features column reports the final number of features with which the model was fitted. A description of the features can be found in Table S1

Dataset	Model	Features	Accuracy	Kappa	AUC	F_1 score	TPR	TNR
1. All variables	SVM	9	0.86 ± 0.1	0.69 ± 0.2	0.91 ± 0.14	0.87 ± 0.12	0.91 ± 0.14	0.82 ± 0.18
2. Significant variables	kNN	7	0.77 ± 0.06	0.52 ± 0.11	0.83 ± 0.12	0.79 ± 0.07	0.78 ± 0.15	0.85 ± 0.16

TABLE 3 Results of the hypothesis contrast approach ML models. Letters in the dataset column refer to the corresponding dataset described in the data analysis section. The features column reports the final number of features with which the model was fitted. A description of the features can be found in Table S1

Dataset	Model	Features	Accuracy	Kappa	AUC	F1	TPR	TNR
(a) Persons/items	SVM	8	0.84 ± 0.13	0.64 ± 0.31	0.88 ± 0.15	0.87 ± 0.11	0.94 ± 0.12	0.71 ± 0.3
(b) Children/ items	Random forest	6	0.82 ± 0.09	0.56 ± 0.3	0.84 ± 0.22	0.86 ± 0.07	0.92 ± 0.12	0.64 ± 0.34
(c) Adults/items	Naïve Bayes	7	0.83 ± 0.14	0.64 ± 0.32	0.84 ± 0.22	0.84 ± 0.12	0.82 ± 0.17	0.85 ± 0.29
(d) Children/adults	SVM	9	0.78 ± 0.18	0.52 ± 0.37	0.79 ± 0.20	0.82 ± 0.16	0.90 ± 0.16	0.61 ± 0.29
(e) Face/body	XGBoost	2	0.73 ± 0.14	0.44 ± 0.26	0.77 ± 0.13	0.77 ± 0.14	0.8 ± 0.16	0.68 ± 0.25
(f) Children's faces/bodies	kNN	8	0.74 ± 0.1	0.38 ± 0.25	0.67 ± 0.25	0.79 ± 0.1	0.88 ± 0.14	0.54 ± 0.28
(g) Adults' faces/bodies	kNN	5	0.66 ± 0.19	0.25 ± 0.36	0.64 ± 0.2	0.7 ± 0.21	0.73 ± 0.19	0.51 ± 0.27
(h) Dynamic/static items	Naïve Bayes	5	0.69 ± 0.17	0.39 ± 0.3	0.72 ± 0.19	0.71 ± 0.17	0.67 ± 0.21	0.71 ± 0.24

seen in the electronics shop; 7. NFrames PersonsChildren_ElectronicShop: number of frames of children seen in the electronics shop; 8. NFramesPersonsChildrenBody_Electronic Shop: number of frames of children's bodies seen in the electronics shop; 9. NFramesPersonsChildrenFace_ElectronicShop: number of frames of children's faces seen in the electronics shop; 10. NFramesPersons Face_ElectronicShop: number of frames of people's faces seen in the electronics shop; 11. NFramesPersonsFace_GameCenter: number of frames of people's faces seen in the game center; 12. NFramesPersonsNoAvatar_Hall: number of frames of people seen in the hall.

DISCUSSION

The ASD gold-standard assessment is based on the manifestation of explicit symptoms through semi-structured observational activities and interviews, in which clinicians attribute mainly qualitative index scores to children's behaviors. However, many ASD dimensions are internal and do not manifest until 2–3 years of age. This delayed manifestation lengthens the time of diagnosis, consequently influencing the possibility of offering early treatments to improve autistic children's functional skills. In accordance with this vision, researchers are attempting to improve methods of diagnosing ASD through the predictive value of behavioral biomarkers (Alcañiz, Chicchi Giglioli, et al., 2020; Alcañiz, Marín-Morales, et al., 2020). The primary aim of this study was to assess virtual social and nonsocial visual information processing and cognition in children with ASD compared with TD children through eye-tracking paradigms as a measure of attunement to, and extraction of, socially relevant information. The second aim was to recognize children with ASD and differentiate them from TD children using ML methods. Specifically, we explored (1) whether it was possible to discriminate between the two populations using eye gaze data and (2) if so, which eye gaze parameters could best distinguish the two populations. The results are discussed in terms of three points: (1) significant differences between groups in eye movements; (2) the performance of ML models in using eye movements to recognize autistic children and features used; and (3) limitations and future studies.

Significant differences between groups in eye movements

The first aim was to identify differences in terms of AOIs with regard to social and nonsocial visual information processing for autistic versus TD children. Figure 4 (boxplots 1 and 2) shows significant differences between the two populations, generally indicating that children with ASD watched social elements (i.e., children's or adults' faces and bodies) in more frames than TD children, both overall and per scenario (in the present study, "scenarios" refer to the different shopping and entertainment stores in which various virtual agents interacted with each participant). This first result is partially in opposition with previous studies on social visual attention in ASD, which have shown mixed and

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

heterogeneous results and have not yet reached a consensus. Specifically, as mentioned in the introduction, some studies have found that autistic children show reduced visual attention to social stimuli (rather than nonsocial stimuli) compared with TD children. In addition, these studies have shown that, when autistic children look at social stimuli, their visual attention is focused more on peripheral areas of the face and/or body and/or the background, rather than eyes and mouth, unlike TD children. In contrast, we know of four studies that have found that autistic children show a visual attention preference for social stimuli (Chawarska et al., 2012; Elsabbagh et al., 2013; Falck-Ytter et al., 2015; Fujisawa et al., 2014). Nevertheless, other studies have found no differences between autistic and TD children in terms of visual attention to social stimuli (e.g., Birmingham et al., 2011; Freeth et al., 2010, 2011; Kuhn et al., 2010; Marsh et al., 2015; Nadig et al., 2010; Parish-Morris et al., 2013). These differences could depend on the social stimuli used, the use of static (i.e., images) or dynamic (i.e., videos or avatars) stimuli, and the level of social content—that is, the quantity of social elements (i.e., number of persons) presented to participants. Low social content refers to the presentation of one static or dynamic person, whereas high social content refers to the presentation of more than one person. Chita-Tegmark's (2016a) recent meta-analysis showed that autistic children spent less time attending to social stimuli than TD children when the social content was higher (i.e., the number of people exceeded one). This result seems to suggest that the higher the amount of social content, the more difficulty autistic children experience monitoring the social environment, thus partially explaining the mixed and heterogeneous results in the previous literature. Contrasting this meta-analysis result, our results showed that rich immersive VR environments characterized by high social content, similar to real environments, seemed to activate more visual behaviors in autistic children than in TD children, generating more eye visualizations of the scenarios. The present result may also rely on the type of device used to present stimuli. Indeed, to our knowledge, no studies have investigated the combined use of head-mounted eye-tracking devices and rich immersive VR environments to assess social and nonsocial visual attention in autistic children. The present study is the first attempt in the field. According to Guillon et al. (2014), findings related to offline social cognition (i.e., the presentation of static and dynamic social stimuli on desktop devices) may not be generalizable across contexts. Therefore, it is plausible that findings in a novel context, such as the present one, might differ from previous results. Procedures related to offline social cognition (Schilbach, 2014) might be far from representing real situations, whereas involving real settings might be critical for autistic children and uncontrollable for the experimenter. Our paradigm involving head-mounted eye-tracking devices and VR seems to represent an effective middle point in this continuum: It gives control over the experimental situation, provides users with realistic and ecologically valid settings, and is suitable for autistic children, who visually prefer realistic situations over real contexts (Cardon & Azuma, 2012).

A second interesting result of our study concerns differences in visual attention behaviors toward nonsocial static stimuli. Specifically, our results showed that autistic children seem to look at static stimuli more than TD children do (Figure 4, boxplots 3 and 4). This result is in line with previous studies' findings that autistic children prefer and spend more time looking at nonsocial static stimuli compared with TD children, suggesting that autistic children show a reduced ability to monitor and manage static and dynamic social content and interactions (Chita-Tegmark, 2016b).

Third, results for autistic children showed a significant higher visual attention preference for adults (Figure 4, boxplot 5) and a significant, moderately higher attention preference for children (the main avatar and other children) compared with TD children (Figure 4, boxplots 6 and 7). These results suggest that, in a dynamic complex VR environment, autistic children generally attend more to social scenes than do TD children—in particular, a significant higher visual preference for adults, who are generally autistic children's habitual interlocutors in relation to their peers. Furthermore, autistic children showed a significant higher preference for looking at people's faces (including both adults and children) in three virtual scenes (Figure 4, boxplots 10–12) compared with TD children. Finally, autistic children showed a visual attention preference for children's bodies and faces compared with TD children (Figure 4, boxplots 8 and 9). No significant results were found regarding adults' bodies and faces. According to the eye-avoidance hypothesis (Tanaka & Sung, 2016), individuals with ASD tend to avoid looking at the eyes in static and dynamic social stimuli due to the discomfort caused by looking at the eye region. Dynamic complex VR environments, however, might reduce this discomfort due to the individual perception of better ability to cope with these types of stimuli, thereby enhancing autistic children's tendency to look at children's faces. Overall, these results suggest that dynamic virtual scenarios with high social content—that is, involving the presentation of more than one person and scene—can elicit unconscious visual behaviors that traditional assessment settings and methodology do not evoke or capture.

Performance of ML models in using eye movements to recognize autistic children and features used

To our knowledge, we have proposed the first supervised ML and eye-tracking paradigm in an immersive VR environment for distinguishing between autistic and TD children in visual social attention behaviors.

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

First, results for the general approach ML models including all variables (9) achieved the best result, with 86% accuracy in the recognition of autistic children (kappa = 0.69) and average sensitivity and specificity of 0.91 and 0.82, respectively. This indicates a balanced model in terms of both conditions predictiveness. This result suggests that the immersive virtual system was able to recognize differences between autistic and TD participants in eye gaze behaviors, highlighting and aligning with previous studies showing atypical visual social attention behaviors in autistic children compared with TD children (Chita-Tegmark, 2016a).

Second, the model that used statistically significant variables also achieved positive results, achieving 77% accuracy in the recognition of autistic children (kappa = 0.52), with a sensitivity of 0.78 and a specificity of 0.85. Both results obtained with this approach reported the lowest standard deviations, showing the highest consistency with respect to the total ML models. This result supports and confirms the statistically significant results in which autistic children showed social attention behaviors that differed from TD children, including higher visual preferences for adults over children as well as specific preferences for adults' faces rather than children's faces, which they looked at more than bodies. This result is partially consistent with previous work. On the one hand, as mentioned in the previous section, the majority of studies using eye-tracking paradigms have shown that children with ASD produced fewer eye movements than TD children. Our results indicated contrary eye gaze behaviors. This may depend on the higher social content that VR can provide, which allowed participants to experience situations similar to real ones. On the other hand, the higher ASD visual social preferences for adults' faces over children's faces and for children's bodies over adults' bodies are consistent with previous studies, and ML approaches could be a valid method with which to overcome the heterogeneity of the previous literature.

Furthermore, our results showed that the model containing information about the number of frames involving people (including adults and children) and items (including static and dynamic stimuli) obtained the best accuracy (84%), with an average sensitivity of 0.94 and average specificity of 0.71. Similar results were achieved with the set of variables containing information about the number of frames of children and items, with 82% accuracy (kappa = 0.56), and the number of frames of adults and items, with 83% accuracy (kappa = 0.64). This last model was more balanced in terms of condition (sensitivity = 0.82; specificity = 0.85). These results suggest that the eye gaze behaviors of autistic children show different patterns than those of TD children with respect to people and items. Furthermore, the results obtained for the model containing information about the number of frames of children versus adults achieved slightly lower accuracy (78%, kappa = 0.52) with lower balance in terms of condition (sensitivity = 0.90; specificity = 0.61), showing that different ASD social attention behaviors can be observed in response to only social stimuli.

Finally, the models related to the number of frames of faces versus bodies (accuracy = 73%, kappa = 0.44), adults and children's faces and bodies (adult face/body accuracy = 66%, kappa = 0.25; child face/body accuracy = 74%, kappa = 0.38), and the number of frames related to dynamic and static items (accuracy = 69%, kappa = 0.39) did not achieve high accuracy or kappa values, indicating that they could not successfully distinguish between autistic and TD children.

Limitations and future studies

This research presented an exploratory analysis of eye gaze as a biomarker of ASD. Although the results are promising, there are also some limitations. First, our sample size is limited, the number of features is high, and the models were not applied to an independent sample. Data from all participants were used to validate the model; therefore, the reported results correspond to the cross-validation process. However, 20 iterations were performed (five folds, four repetitions) and low standard deviations were reported to support the extrapolability of the results. Moreover, the models were built using the default hyperparameters so as not to increase the complexity of the problem by attempting to tune them. Future studies on larger samples should allow the model to be tested, improving the generalization of the results. Second, ML models have a high computational cost. Training all of the models presented in this paper took about 20 hours, even when parallelized. However, once trained, the models could be used on simpler computers with computation times of around seconds. Third, experimental groups were not matched on sociodemographics, IQ, or cognitive ability, limiting the generalization of the model. Future studies should consider these features as possible moderators of the development of ASD.

Fourth, the AOI approach used here provided significant information capable of distinguishing autistic from TD children, but mainly referred to macro-areas, such as people (including adults and children), items (including static and dynamic stimuli), and adults and children versus items. However, our results indicate that this approach presented some limitations in terms of recognition of specific eye gaze patterns related to specific areas, such as face areas (i.e., eyes and mouth). Future studies should refine the definition and configuration of more specific AOIs, including eye and mouth AOIs, and consider a bottom-up or data-driven approach in order to include more variables with respect to real eye gaze patterns.

Furthermore, another possible limitation at the technological level may result from the use of eye-tracking glasses that may not be accepted by all children, especially those who are particularly young, thus negatively influencing the likelihood of early diagnosis of autism. The cost of such

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

devices is also very high (as is the assembly of a CAVE), which constitutes an important limitation in the widespread use of the technological system by clinicians. For these reasons, the development of a portable technological system with an integrated remote eye-tracking system is being considered for future projects and studies.

Finally, future studies that include other neurodevelopmental disorders involving social and communication impairments, such as attention-deficit/hyperactivity disorder, could improve discrimination between groups compared with TD children and identify with greater precision the degrees of severity in each disorder, highlighting patterns, similarities, and differences.

CONCLUSIONS

Social attentional cue abilities—including joint attention, paralinguistic head and body movements, and linguistic verbalizations—are related to some of the first symptoms to manifest in autistic children. Traditional ASD assessments based on clinicians' expert evaluations through semi-structured observational tasks in laboratory settings are not able to objectively capture children's internal dimensions or behaviors in real-life situations. The combination of VR, behavioral biomarkers, and ML techniques can provide earlier diagnoses, as well as more objective and precise evaluations in more ecologically valid situations. This, when combined with traditional measures, can enhance knowledge on the internal and implicit dimensions of ASD as well as the development of tailored treatments. In this framework, the current study has shown a proof of concept for diagnosis via the use of a possible disruptive method to assess autistic children at an earlier stage.

ACKNOWLEDGMENTS

This work was supported by the Spanish Ministry of Economy, Industry, and Competitiveness-funded project "Immersive Virtual Environment for the Evaluation and Training of Children with Autism Spectrum Disorder: T Room" (IDI-20170912).

ORCID

Irene Alice Chicchi-Giglioli https://orcid.org/0000-0003-2577-0039

REFERENCES

Alcañiz Raya, M., Chicchi Giglioli, I. A., Marín-Morales, J., Higuera-Trujillo, J. L., Olmos, E., Minissi, M. E., Teruel-García G., Sirera, M., & Abad, L. (2020). Application of supervised machine learning for behavioral biomarkers of autism spectrum disorder based on electrodermal activity and virtual reality. *Frontiers in Human Neuroscience*, *14*, 90.

Alcañiz Raya, M., Marín-Morales, J., Minissi, M. E., Teruel Garcia, G., Abad, L., & Chicchi Giglioli, I. A. (2020). Machine learning and virtual reality on body movements' behaviors to classify children with autism spectrum disorder. *Journal of Clinical Medicine*, *9*(5), 1260.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

Birmingham, E., Cerf, M., & Adolphs, R. (2011). Comparing social attention in autism and amygdala lesions: Effects of stimulus and task condition. *Social Neuroscience*, *6*(5–6), 420–435.

Bischl, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., Casalicchio, G., & Jones, Z. (2016). Mlr: Machine learning in R. *Journal of Machine Learning Research*, *17*(170), 1–5. http://jmlr.org/papers/v17/15-066.html

Bridgemohan, C., Cochran, D. M., Howe, Y. J., Pawlowski, K., Zimmerman, A. W., Anderson, G. M., … Neumeyer, A. M. (2019). Investigating potential biomarkers in autism spectrum disorder. *Frontiers in Integrative Neuroscience*, *13*, 31. https://doi.org/10.3389/fnint.2019.00031

Burdea, G. C., & Coiffet, P. (2003). Virtual reality technology. In *Virtual reality technology*. John Wiley & Sons.

Cai, Y., Chia, N. K., Thalmann, D., Kee, N. K., Zheng, J., & Thalmann, N. M. (2013). Design and development of a virtual dolphinarium for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*(2), 208–217.

Cardon, T., & Azuma, T. (2012). Visual attending preferences in children with autism spectrum disorders: A comparison between live and video presentation modes. *Research in Autism Spectrum Disorders*, *6*(3), 1061–1067.

Carette, R., Cilia, F., Dequen, G., Bosche, J., Guerin, J. L., & Vandromme, L. (2017). Automatic autism spectrum disorder detection thanks to eye-tracking and neural network-based approach. *International Conference on IoT Technologies for HealthCare*, 75–81.

Carette, R., Elbattah, M., Dequen, G., Guérin, J., Cilia, F., & Bosche, J. (2019). Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths. *Proceedings of the 12th International Conference on Health Informatics*, pp. 103–112.

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry*, *53*(8), 903–913.

Chevallier, C., Parish-Morris, J., Mcvey, A., Rump, K., Noah, J., Herrington, J., et al. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, *8*, 620–628. http://doi.org/10.1002/aur.1479

Chita-Tegmark, M. (2016a). Attention allocation in ASD: A review and meta-analysis of eye-tracking studies. *Review Journal of Autism and Developmental Disorders*, *3*(3), 209–223.

Chita-Tegmark, M. (2016b). Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Research in Developmental Disabilities*, *48*, 79–93.

Cilia, F., Aubry, A., Le Driant, B., Bourdin, B., & Vandromme, L. (2019). Visual exploration of dynamic or static joint attention bids in children with autism syndrome disorder. *Frontiers in Psychology*, *10*.

Cipresso, P., Giglioli, I. A. C., Raya, M. A., & Riva, G. (2018). The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature. *Frontiers in Psychology*, *9*, 2086.

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.

Del, M., Rubido, V., Mccracken, J. T., Hollander, E., Shic, F., Noeldeke, J., … Umbricht, D. (2018). In search of biomarkers for autism spectrum disorder. *Autism Research*, *11*, 1567–1579. https://doi.org/10.1002/aur.2026

Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., Johnson, M. H., & BASIS Team. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behavioural Brain Research*, *251*, 147–154.

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Falck-Ytter, T., Carlström, C., & Johansson, M. (2015). Eye contact modulates cognitive processing differently in children with autism. *Child Development*, *86*(1), 37–47.

Freeth, M., Ropar, D., Chapman, P., & Mitchell, P. (2010). The eye gaze direction of an observed person can bias perception, memory, and attention in adolescents with and without autism spectrum disorder. *Journal of Experimental Child Psychology*, *105*(1–2), 20–37.

Freeth, M., Ropar, D., Mitchell, P., Chapman, P., & Loher, S. (2011). Brief report: How adolescents with ASD process social information in complex scenes. Combining evidence from eye movements and verbal descriptions. *Journal of Autism and Developmental Disorders*, *41*(3), 364–371.

Fujisawa, T. X., Tanaka, S., Saito, D. N., Kosaka, H., & Tomoda, A. (2014). Visual attention for social information and salivary oxytocin levels in preschool children with autism spectrum disorders: An eye-tracking study. *Frontiers in Neuroscience*, *8*, 295.

Gonçalves, N., Rodrigues, J. L., Costa, S., & Soares, F. (2012). Preliminary study on determining stereotypical motor movements. *2012 Annual International Conference of The IEEE Engineering in Medicine and Biology Society*, 1598–1601.

Guazzaroni, G. (Ed.). (2018). *Virtual and augmented reality in mental health treatment*. IGI Global.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience & Biobehavioral Reviews*, *42*, 279–297.

He, Q., Wang, Q., Wu, Y., Yi, L., & Wei, K. (2021). Automatic classification of children with autism spectrum disorder by using a computerized visual-orienting task. *PsyCh Journal.*, *10*, 550–565.

Kuhn, G., Kourkoulou, A., & Leekam, S. R. (2010). How magic changes our expectations about autism. *Psychological Science*, *21*(10), 1487–1493.

Liu, W., Li, M., & Yi, L. (2016). Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. *Autism Research*, *9*(8), 888–898.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2001). *Autism diagnostic observation schedule*. Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. A. (1999). *Autism diagnostic observation schedule-WPS (ADOS-WPS)*. Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Marsh, L. E., Pearson, A., Ropar, D., & Hamilton, A. D. C. (2015). Predictive gaze during observation of irrational actions in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *45*(1), 245–261.

Min, C. H., & Tewfik, A. H. (2010). Automatic characterization and detection of behavioral patterns using linear predictive coding of accelerometer sensor data. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 220–223.

Minissi, M. E., Chicchi Giglioli, I. A., Mantovani, F., & Raya, M. A. (2021). Assessment of the autism spectrum disorder based on machine learning and social visual attention: A systematic review. *Journal of Autism and Developmental Disorders*, 1–16.

Nadig, A., Lee, I., Singh, L., Bosshart, K., & Ozonoff, S. (2010). How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism. *Neuropsychologia*, *48*(9), 2730–2739.

Parish-Morris, J., Chevallier, C., Tonge, N., Letzen, J., Pandey, J., & Schultz, R. T. (2013). Visual attention to dynamic faces and objects is linked to face processing skills: A combined study of children with autism and controls. *Frontiers in Psychology*, *4*, 185.

Parsons, S. (2016). Authenticity in virtual reality for assessment and intervention in autism: A conceptual review. *Educational Research Review*, *19*, 138–157.

Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research*, *46*(5), 430–443. https://doi.org/10.1046/j.1365-2788.2002.00425.x

Ruggeri, B., Sarkans, U., Schumann, G., & Persico, A. M. (2014). Biomarkers in autism spectrum disorder: The old and the new. *Psychopharmacology*, *231*(6), 1201–1216. https://doi.org/10.1007/s00213-013-3290-7

Rutherford, M. D., & Towns, A. M. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(7), 1371–1381.

Saitovitch, A., Bargiacchi, A., Chabane, N., Phillipe, A., Brunelle, F., Boddaert, N., … Zilbovicius, M. (2013). Studying gaze abnormalities in autism: Which type of stimulus to use?. *Open Journal of Psychiatry*, *3*, 32–38.

Schaller, U. M., Biscaldi, M., Burkhardt, A., Fleischhaker, C., Herbert, M., Isringhausen, A., & Rauh, R. (2021). ADOS-Eye-Tracking: The archimedean point of view and its absence in Autism Spectrum Conditions. *Frontiers in Psychology*, *12*, 497.

Schilbach, L. (2014). On the relationship of online and offline social cognition. *Frontiers in Human Neuroscience*, *8*, 278.

Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry*, *75*, 231–237. https://doi.org/10.1016/j.biopsych.2013.07.009

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557.

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism*, *11*(3), 265–277.

Sumiya, M., Okamoto, Y., Koike, T., Tanigawa, T., Okazawa, H., Kosaka, H., & Sadato, N. (2020). Attenuated activation of the anterior rostral medial prefrontal cortex on self-relevant social reward processing in individuals with autism spectrum disorder. *NeuroImage: Clinical*, *102249*.

Tanaka, J. W., & Sung, A. (2016). The "eye avoidance" hypothesis of autism face processing. *Journal of Autism and Developmental Disorders*, *46*(5), 1538–1552.

Thorup, E., Nyström, P., Gredebäck, G., Bölte, S., & Falck-Ytter, T. (2016). Altered gaze following during live interaction in infants at risk for autism: An eye tracking study. *Molecular Autism*, *7*(1), 1–10.

van der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N., & van Engeland, H. (2002). Looking at images with human figures: Comparison between autistic and normal children. *Journal of Autism and Developmental Disorders*, *32*(2), 69–75.

Vettori, S., Dzhelyova, M., Van der Donck, S., Jacques, C., Steyaert, J., Rossion, B., & Boets, B. (2020). Frequency-tagging electroencephalography of superimposed social and non-social visual stimulation streams reveals reduced saliency of faces in autism spectrum disorder. *Frontiers in Psychiatry*, *11*, 332.

Vettori, S., Dzhelyova, M., Van der Donck, S., Jacques, C., Van Wesemael, T., Steyaert, J., … Boets, B. (2020). Combined frequency-tagging EEG and eye tracking reveal reduced social bias in boys with autism spectrum disorder. *Cortex*, *125*, 135–148.

Wan, G., Kong, X., Sun, B., Yu, S., Tu, Y., Park, J., … Kong, J. (2019). Applying eye tracking to identify autism spectrum disorder in children. *Journal of Autism and Developmental Disorders*, *49*(1), 209–215.

Wallace, S., Parsons, S., Westbury, A., White, K., White, K., & Bailey, A. (2010). Sense of presence and atypical social judgments in immersive virtual environments: Responses of adolescents with Autism Spectrum Disorders. *Autism*, *14*(3), 199–213.

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

World Health Organizationhttps://www.who.int/news-room/fact-sheets/detail/autismspectrum-disorders. 2019.

Zhao, Z., Tang, H., Zhang, X., Qu, X., Hu, X., & Lu, J. (2021). Classification of children with autism and typical development using eye-tracking data from face-to-face conversations: Machine learning model development and performance evaluation. *Journal of Medical Internet Research*, *23*(8), e29328.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Alcañiz, M., Chicchi-Giglioli, I. A., Carrasco-Ribelles, L. A., Marín-Morales, J., Minissi, M. E., Teruel-García, G., Sirera, M., & Abad, L. (2022). Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis. *Autism Research*, *15*(1), 131–145. https://doi.org/10.1002/aur.2636

19393806, 2022, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aur.2636 by Universite Du Quebec Abitibi, Wiley Online Library on [05/02/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License