Aus der
Klinik und Poliklinik für Psychiatrie und Psychotherapie
Klinikum der Ludwig-Maximilians-Universität München

Perception and production of non-verbal signals in autism

Dissertation
zum Erwerb des Doctor of Philosophy (Ph.D.)
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von
Marta Robles Rosa

aus
Granollers / Spanien

Jahr
2025

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten:	Prof. Dr. Christine Falter-Wagner
Zweites Gutachten:	Prof. Dr. Zhuanghua Shi
Drittes Gutachten:	Prof. Dr. Sophia Stöcklein
Viertes Gutachten:	Prof. Dr. Anne Mareike Altgassen

Dekan:	Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung:	21.02.2025

Affidavit

LMU

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Promotionsbüro
Medizinische Fakultät

MMRS

Affidavit

Robles Rosa, Marta
____
Surname, first name

Nußbaumstraße, 7
____
Street

80336, Munich, Germany
____
Zip code, town, country

I hereby declare, that the submitted thesis entitled:

Perception and production of non-verbal signals in autism

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree, except as provided in the respective Cotutelle Agreement.

Munich, 24th of September 2024
____
place, date

Marta Robles Rosa
____
Signature doctoral candidate

Confirmation of congruency

LMU

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Promotionsbüro
Medizinische Fakultät

MMRS

Confirmation of congruency between printed and electronic version of the doctoral thesis

Robles Rosa, Marta

Surname, first name

Nußbaumstraße, 7

Street

80336, Munich, Germany

Zip code, town, country

I hereby declare, that the submitted thesis entitled:

Perception and production of non-verbal signals in autism

is congruent with the printed version both in content and format.

Munich, 24th September 2024

place, date

Marta Robles Rosa

Signature doctoral candidate

Ethical commitment statement

This thesis adhered to the principles of the Declaration of Helsinki (World Medical Association [WMA], 2013). Additionally, the included experiments were evaluated and approved by the Research Ethics Committee (CERec) of the Autonomous University of Barcelona (UAB) (Publication I) and by the Ethics Committee of the Medical Faculty of the LMU Munich (LMU) (Publication II).

All participants were provided with comprehensive information regarding the study, including the aim and methods of the study, as well as their right to refuse participation or withdraw consent at any time without reprisal. Informed consent was obtained from all participants prior to commencing the study.

Additionally, I affirm that there are no conflicts of economic, ideological, or any other nature that could potentially influence this doctoral thesis.

Presentation

This thesis explored non-verbal signals in autism. To achieve this, two independent studies have been conducted and are presented within a publication-based thesis. The first article (hereafter Publication I), "*Reduced stereotypicality and spared use of facial expression predictions for social evaluation in autism*" was published in 2024 in the *International Journal of Clinical and Health Psychology*, explored how autistic individuals perceive social stimuli, more specifically, how they make predictions about facial expressions, and use these predictions for social evaluation. The second article (henceforth Publication II), "*A Virtual Reality Based System for the Screening and Classification of Autism*", was published in 2022 in the journal *IEEE Transactions on Visualization and Computer Graphics* and presents a feasibility or proof-of-concept study exploring the potential application of an immersive virtual reality (VR) system designed to objectively measure eye movements and other non-verbal communication features in autism, also assessing the use of this data for machine learning (ML) classification applications. Both publications are open-access.

This thesis has been conducted under a Cotutelle Agreement between the LMU and the UAB. Within the framework of this agreement, Publication I was conducted at the UAB, under the main supervision of Professor Dr. Lorena Chanes Puiggros, and Publication II was conducted at the LMU under the main supervision of Professor Dr. Christine M. Falter-Wagner.

This thesis is structured in different sections and represents the most viable arrangement to meet the regulatory requirements of both universities involved in the cotutelle agreement. Firstly, an introduction serves as the background to contextualise the investigation and my contribution to this thesis. It includes a general explanation of autism, along with subsections focusing on non-verbal signals in autism, revising the state of the art, and identifying research opportunities that this thesis tackled. Next, the aims and hypotheses of the thesis are outlined, providing a comprehensive overview of the research objectives. This is followed by short summaries of Publication I and Publication II. Subsequently, a discussion is presented, synthesising the findings from both studies and drawing overarching conclusions. Lastly, a list of references is included, ensuring acknowledgement of all sources consulted and cited throughout the thesis. Full copies of the publications are also provided, depicting each publication's methodology, results, and independent discussions thoroughly. The thesis also includes an ethical commitment statement, a note on language terminology, and a list of publications detailing my personal contributions, among other relevant sections.

Table of content

Note on terminology

In recent years, a contentious discussion has emerged within the autism community regarding the optimal terminology for describing individuals who received a diagnosis of autism. While different terms are used to refer to a person with such diagnosis, the essence of the debate lies in employing person-first language (i.e., 'person with autism') or embracing identity-first language (i.e., 'autistic person'). In the past decade, this topic has gained prominence within the scientific community, with numerous studies worldwide examining the perspectives of various stakeholders in the autism community, including autistic (Bosman & Thijs, 2024; Botha et al., 2022; Buijsman et al., 2023; Bury et al., 2020, 2023; De Laet et al., 2023; Geelhand et al., 2023; Kapp et al., 2013; Keating et al., 2023; Kenny et al., 2015; Lei et al., 2021; Shakes & Cashin, 2019; Taboas et al., 2022). Despite the lack of consensus, the investigations offer insights into why this situation may be occurring. For instance, studies show that different English-speaking stakeholders have different preferences: while most autistic and self-identified adults prefer identity-first terminology, parents and professionals seem to prefer employing person-first language (Kapp et al., 2013; Kenny et al., 2015; Taboas et al., 2022). Across national boundaries, spoken language appears to influence perceptions, as evidenced by a study conducted in Belgium involving Dutch-speaking adults with autism (De Laet et al., 2023), which contradicted findings from a separate study also in Belgium involving French-speaking adults with autism (Geelhand et al., 2023). The former study (De Laet et al., 2023) aligns with previous research (Bosman & Thijs, 2024; Buijsman et al., 2023) indicating that the majority of Dutch-speaking adults with autism prefer person-first terms. On the contrary, the latter (Geelhand et al., 2023) concluded that the majority of French-speaking adults with autism also living in Belgium preferred identity-first language. Other studies show that individuals with a stronger sense of autism identity tend to favour identity-first terms (Bosman & Thijs, 2024; Bury et al., 2023), as well as higher awareness of the neurodiversity movement (Kapp et al., 2013). Moreover, findings by Bury et al. (2023) suggest that individuals experiencing less stigma and lower levels of internalised stigma tend to prefer identity-first language.

Studies of this nature are presently lacking in German- and Spanish-speaking countries, which correspond to the locations of the studies included in this thesis. Nevertheless, I have chosen to adopt the terminology that aligns with the overall reported preference of most adult autistic individuals, which is identity-first language (Botha et al., 2022; Bury et al., 2020; Geelhand et al., 2023; Kapp et al., 2013; Keating et al., 2023; Kenny et al., 2015; Lei et al., 2021; Taboas et al., 2022).

However, since views on terminology are highly personal (Monk et al., 2022; Tepest, 2021), I would like to acknowledge the coexistence of different language preferences in referring to a person with a diagnosis of autism, emphasising the importance of unequivocally respecting the preferences of minority groups.

Finally, on that note, for better readability as well as unambiguity, I will refer to the comparison groups as “control” groups. In the case of the investigations included in this thesis, this term refers to samples of people who reported absence and no history of any psychiatric or neurological conditions, including autism, as well as no intake of psychoactive medication. Other studies and literature may refer to these groups as the “typically developed” group, “neurotypical” group, or “non-autistic” group.

List of abbreviations

ADI-R: Autism Diagnostic Interview-Revised

ADOS-2: The Autism Diagnostic Observation Schedule, Second Edition

AWMF: Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften

APA: American Psychiatric Association

CDC: Centers for Disease Control and Prevention

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

DSM-5-TR: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revised

ESCAP: European Society of Child and Adolescent Psychiatry

fMRI: Functional Magnetic Resonance Imaging

H: Hypothesis

VR: Virtual Reality

ML: Machine Learning

NHS: National Health Service

NICE: National Institute for Health and Care Excellence

NIMH: National Institute of Mental Health

RQ: Research Question

WHO: World Health Organization

WMA: World Medical Association

List of publications and author contributions

This thesis is based on the following published original studies:

- Publication I:

Robles, M., Ramos-Grille, I., Hervás, A., Duran-Tauleria, E., Galiano-Landeira, J., Wormwood, J.B., Falter-Wagner, C.M., & Chanes, L. (2024). Reduced stereotypicality and spared use of facial expression predictions for social evaluation in autism. *International Journal of Clinical and Health Psychology*, *24*(2), 100440. https://doi.org/10.1016/j.ijchp.2024.100440

Journal Citation Reports: International Journal of Clinical and Health Psychology

Impact Factor 2022: 8.8

Journal Area: Psychology

Ranked 5th of 131 clinical psychology journals (Q1)

Personal contribution to Publication I: I contributed to preparing the Institutional Review Board ethics amendment for the study. I assumed responsibility for study administration, including finding and establishing a collaboration with a recruitment centre, management and coordination for the research activity, planning and execution. I was very much involved in the investigation, including participants' recruitment, and data collection. I was responsible for data curation, metadata annotation, data pre-processing, and maintaining research data for initial and later use. I performed all formal analyses and mainly prepared the figures and table visualizations. I contributed to the original draft writing, including preparation, creation, and presentation of the published work for its publication and dissemination at national and international conferences.

- Publication II:

Robles, M.*, Namdarian, N.*, Otto, J.*, Wassiljew, E., Navab, N., Falter-Wagner, C. M., & Roth, D. (2022). A Virtual Reality Based System for the Screening and Classification of Autism. *IEEE Transactions on Visualization and Computer Graphics, 28*(5), 2168–2178. https://doi.org/10.1109/TVCG.2022.3150489

*Shared first authorship.

Journal Citation Reports: IEEE Transactions on Visualization and Computer Graphics

Impact Factor 2022: 5.2

Journal Area: Computer Science

Ranked 15th of 108 Computer Science journals (Q1)

Personal contribution to Publication II: I contributed to the design of the study's methodology, partially designing and producing the content of the VR task, preparing instructions to support the explanation of the task, as well as selecting the most appropriate questionnaires to use. I also mainly prepared the Institutional Review Board ethics application and subsequent amendments, and I assumed responsibility for study administration, including management and coordination for the research activity, planning and execution. I was significantly involved in the investigation, including participants' recruitment, experiment set-up, and data collection. I was responsible for data curation, metadata annotation, psychological data processing, and maintaining research data for initial and later use. I contributed mainly to group comparison analyses and their visualization, creating tables and graphs, as well as interpretation of results. I contributed to the original draft writing, including preparation, creation and presentation of the published work for its publication, and dissemination at national and international conferences.

Additional publications not included in this work:

- Alonzo-Castillo, T., Lugo-Marín, J., **Robles, M.**, Rossich, R., Gallego, L., González, M., Setién-Ramos, I., Martínez-Ramírez, M., Ramos-Quiroga, J. A., & Gisbert-Gustemps, L. (2024). Trastorno del espectro autista: impacto de una estrategia de formación en línea en los conocimientos del personal sanitario de un hospital de tercer nivel [Autism spectrum disorder: the impact of an online training strategy on the knowledge of the healthcare staff of a tertiary care hospital]. *Revista de Neurologia*, *78*(1), 1–7. https://doi.org/10.33588/rn.7801.2023244

- Draganov, M., Galiano-Landeira, J., Doruk Camsari, D., Ramírez, J. E., **Robles, M.**, & Chanes, L. (2023). Noninvasive modulation of predictive coding in humans: causal evidence for frequency-specific temporal dynamics. *Cerebral cortex*, *33*(13), 8421–8430. https://doi.org/10.1093/cercor/bhad127

- Falter-Wagner, C. M., Bloch, C., **Robles, M.**, Horch, L., Vogeley, K., & Georgescu, A. L. (2022). Figure-Disembedding Is Inferior in Non-autistic Compared to Autistic Individuals but Can Be Improved by Training. *Frontiers in Psychology*, *13*, 857630. https://doi.org/10.3389/fpsyg.2022.857630

- Ramos-Grille, I., Weyant, J., Wormwood, J. B., **Robles, M.**, Vallès, V., Camprodon, J. A., & Chanes, L. (2022). Predictive processing in depression: Increased prediction error following negative valence contexts and influence of recent mood-congruent yet irrelevant experiences. *Journal of Affective Disorders*, *311*, 8–16. https://doi.org/10.1016/j.jad.2022.05.030

Abstract

Background: Autism is a neurodevelopmental condition including characteristic communication and social interaction difficulties, and restricted, repetitive patterns of behaviour, interests, or activities. Within the communication domain, differences in non-verbal signals, such as facial expressions or eye contact during a social interaction, have been reported in autistic individuals with regards to control individuals.

Objectives: The general aim of this thesis was to explore non-verbal signals in autism, particularly, facial expressions and eye, head and hand movements. The specific objectives were (1) to explore how autistic individuals make predictions about facial expressions and how these predictions are used for social evaluation, and (2) to explore the potential of eye, head and hand movements as markers for autism in a virtual reality (VR) setting, to further use these data for pattern classification.

Methods: Two studies were conducted. In Publication I, adult autistic individuals (n = 34) and adult control individuals (n = 34) were assessed using a social perception behavioural paradigm to explore facial expression predictions and their impact on social evaluation. In Publication II, eye, head, and hand movements were recorded in adult autistic individuals (n = 6) and adult control individuals (n = 13) during a VR shopping experience task, interacting with a virtual character acting as the supermarket assistant. This data was then analysed using machine learning techniques for pattern classification.

Results: In Publication I, autistic individuals held less stereotypical predictions about facial expressions than controls. The use of such predictions for social evaluation (likability) did not differ significantly between groups, with autistic individuals relying on their predictions to evaluate others to a similar extent than controls. In Publication II, Significant differences between autistic and control individuals were observed in gaze fixation times and head movements during social interactions. Compared to control, autistic individuals showed reduced fixation times in the eye region of the virtual character and, instead, looked more to the background during social interactions. Utilising a VR system to capture non-verbal communication signals in autism for pattern classification is feasible.

Conclusions: Autistic individuals present significant differences in perception and production of non-verbal signals. Non-verbal communication signals can be objectively captured in a VR-system with strong potential for pattern classification.

Keywords: autism, non-verbal signals, facial expressions, emotion, virtual reality, machine learning

Resumen

Antecedentes: El autismo es una condición del neurodesarrollo que incluye dificultades características en la comunicación y la interacción social, y patrones restringidos y repetitivos de comportamiento, intereses o actividades. Dentro del dominio de la comunicación, se han documentado diferencias en las señales no verbales, como las expresiones faciales o el contacto visual durante una interacción social, en individuos autistas en comparación con individuos controles.

Objetivos: El objetivo general de esta tesis fue explorar las señales no verbales en el autismo, particularmente, las expresiones faciales y los movimientos de los ojos, cabeza y manos. Los objetivos específicos fueron (1) explorar cómo las personas autistas hacen predicciones sobre las expresiones faciales y cómo estas predicciones se utilizan para la evaluación social, y (2) explorar el potencial de los movimientos de los ojos, cabeza y manos como marcadores de autismo en un entorno de realidad virtual (VR), para utilizar estos datos en la clasificación de patrones.

Métodos: Se realizaron dos estudios. En la Publicación I, se evaluaron individuos autistas adultos (n = 34) e individuos controles adultos (n = 34) utilizando un paradigma conductual de percepción social para explorar las predicciones sobre expresiones faciales y su impacto en la evaluación social. En la Publicación II, se registraron los movimientos de los ojos, cabeza y manos de individuos autistas adultos (n = 6) e individuos controles adultos (n = 13) durante una tarea en un entorno de VR, que consistía en hacer una compra de supermercado, interactuando con un personaje virtual que hacía de vendedor. Estos datos se analizaron utilizando técnicas de aprendizaje automático para la clasificación de patrones.

Resultados: En la Publicación I, los individuos autistas tuvieron unas predicciones sobre las expresiones faciales menos estereotípicas en comparación con los individuos controles. El uso de dichas predicciones para la evaluación social (agradabilidad) no difirió significativamente entre los grupos, es decir, los individuos autistas usaron sus predicciones para evaluar a los demás en la misma medida que lo hicieron los controles. En la Publicación II, se observaron diferencias significativas entre individuos autistas y controles en los tiempos de fijación de la mirada y en los movimientos de la cabeza durante las interacciones sociales. En comparación con los controles, los individuos autistas mostraron tiempos de fijación reducidos en la región de los ojos del personaje virtual y, en su lugar, miraron más hacia el fondo durante esas interacciones sociales con el vendedor. Utilizar un sistema de VR para capturar señales de comunicación no verbal en el autismo y para la clasificación de patrones es factible.

Conclusiones: Los individuos autistas presentan diferencias en la percepción y producción de señales de comunicación no verbal. Las señales de comunicación no verbal pueden capturarse objetivamente en un sistema de VR y muestran un gran potencial para la clasificación de patrones.

Palabras clave: autismo, señales no verbales, expresiones faciales, emoción, realidad virtual, aprendizaje automático

Abstrakt

Hintergrund: Autismus ist eine Störung der neuronalen Entwicklung, die durch charakteristische Schwierigkeiten in der Kommunikation und sozialen Interaktion sowie durch eingeschränkte, repetitive Verhaltensmuster, Interessen oder Aktivitäten gekennzeichnet ist. Im Bereich der Kommunikation wurden bei autistischen

Personen Unterschiede in non-verbalen Signalen, wie im Gesichtsausdruck oder im Blickkontakt während sozialer Interaktionen, im Vergleich zu Kontrollpersonen festgestellt.

Ziele: Das zentrale Ziel der vorliegenden Dissertation war es, non-verbale Signale bei autistische Personen zu untersuchen. Der Schwerpunkt lag hierbei auf Gesichtsausdrücken, sowie Augen-, Kopf- und Handbewegungen. Die spezifischen Ziele waren (1) zu erforschen, wie autistische Personen Vorhersagen über Gesichtsausdrücke treffen und wie diese Vorhersagen für soziale Bewertungen verwendet werden, und (2) das Potenzial von Augen-, Kopf- und Handbewegungen als Marker für Autismus in einer virtuellen Realität (VR)-Umgebung zu untersuchen, um diese Daten für Musterklassifikationen zu verwenden.

Methoden: Es wurden zwei Studien durchgeführt. In Publikation I wurden Daten mittels eines Verhaltensexperiments zur sozialen Wahrnehmung von erwachsenen autistischen Personen (n = 34) und erwachsenen Kontrollpersonen (n = 34) erhoben, um Vorhersagen über Gesichtsausdrücke und deren Einfluss auf soziale Bewertungen zu untersuchen. In Publikation II wurden Augen-, Kopf- und Handbewegungen bei erwachsenen autistischen Personen (n = 6) und Kontrollpersonen (n = 13) während einer VR-Einkaufserfahrung aufgezeichnet. Versuchspersonen interagierten hierbei mit einem virtuellen Charakter, welcher die Funktion eines Supermarktmitarbeiters inne hatte. Diese Daten wurden anschließend mit maschinellen Lerntechniken zur Musterklassifikation analysiert.

Ergebnisse: In Publikation I hatten autistische Personen weniger stereotypische Vorhersagen über Gesichtsausdrücke als Personen der Kontrollgruppe. Die Verwendung solcher Vorhersagen zur sozialen Bewertung (Sympathie) unterschied sich nicht signifikant zwischen den Gruppen, wobei autistische Personen in ähnlichem Maße wie Personen der Kontrollgruppe auf ihre Vorhersagen zur Bewertung anderer zurückgriffen. In Publikation II wurden signifikante Unterschiede zwischen autistischen und Kontrollpersonen in Bezug auf Blickfixierungszeiten und Kopfbewegungen während sozialer Interaktionen festgestellt. Im Vergleich zu Kontrollpersonen zeigten autistische Personen reduzierte Fixierungszeiten auf den Augenbereich des virtuellen Charakters. Häufiger schauten sie stattdessen während sozialen Interaktionen in den Hintergrund. Die Nutzung eines VR-Systems zur Erfassung non-verbaler Kommunikationssignale im Autismus zur Musterklassifikation zeigt sich als umsetzbar.

Schlussfolgerungen: Autistische Personen weisen signifikante Unterschiede in der Wahrnehmung und Produktion non-verbaler Signale auf. Non-verbale Kommunikationssignale können objektiv in einem VR-System erfasst werden und zeigen ein großes Potenzial für die Musterklassifikation.

Schlüsselwörter: Autismus, non-verbale Signale, Gesichtsausdrücke, Emotion, virtuelle Realität, maschinelles Lernen

1. Introduction

Autism is a neurodevelopmental condition (or disorder, from the conventional medical-model paradigm, Pellicano & den Houting, 2022; Ritvo & Freeman, 1984), defined by a set of characteristics (or symptoms) persisting throughout life, including (1) difficulties (or deficits) in social communication and reciprocal social interaction and (2) restricted, repetitive patterns of behaviour, interests, or activities (adapted from the text revision of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders [DSM-5-TR], American Psychiatric Association [APA], 2022). First described in 1925 by the psychiatrist and researcher Grunya Efimovna Sukhareva (Sukhareva, 1925), autism prevalence rates have been rising, ranging from 1% (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften [AWMF], 2015) to almost 3% (Centers for Disease Control and Prevention [CDC], Maenner et al., 2023).

Lorna Wing (1981) introduced the concept of an autism “spectrum” (Happé & Baron-Cohen, 2014), describing the significant variation in the type and severity of symptoms experienced by autistic people (National Institute of Mental Health [NIMH], 2024). For example, some autistic individuals present absence—or even avoidance—of eye contact during a social interaction, while other autistic individuals do not present absence but rather reduced—or atypical—eye contact (e.g., reduced modulation of the eye contact) during a social interaction (APA, 2022). As a second example, literature shows that autistic individuals present differences in sensory processing, namely, hypersensitivity (over-reacting) or hyposensitivity (under-reacting) to sensory input (Robertson & Baron-Cohen, 2017), but not all autistic individuals exhibit such characteristics (Milton, 2016). Moreover, while the core distinct features of autism, according to APA (2022), primarily involve the two previously stated dimensions, autistic individuals often have difficulties in the executive functions (Demetriou et al., 2018), language (Friedman & Sterling, 2019), motor abilities (Travers et al., 2013), and social cognition (Lai et al., 2014), among others. This variability in autism can be attributed to several factors, including individual differences in personality (Di Vara et al., 2024), sex and gender (Lai & Szatmari, 2020), varying levels of social skills or intellectual abilities (Saure et al., 2023), timing of diagnosis and intervention (Lai & Baron-Cohen, 2015), and the use of learnt strategies either through experience or the development of coping mechanisms (Ghanouni & Quirke, 2023; Hull et al., 2020). Despite the variability in how autism presents (Wozniak et al., 2017), one of the most notable characteristics of autistic individuals lies in non-verbal communication signals.

Non-verbal signals offer significant opportunities for investigation in autism and are the centrepiece of this thesis. Therefore, these aspects will be examined in more detail.

Non-verbal signals in autism

Non-verbal signals and behaviours are central to socially meaningful outcomes of communicative interactions across all types of social relationships (Manusov & Patterson, 2006). Non-verbal communicative behaviours indicate basic social orientations in how we interact with others and reflect our understanding of social situations (Piaget, 1960). These include all communicative acts except speech (Mandal, 2014). Visible non-verbal modalities extend beyond facial expressions to include head, body, and hand movements, posture, interpersonal gaze, orientation, distance, and interpersonal synchrony or mimicry (Hall & Knapp, 2013). Auditory non-verbal cues consist of discrete vocal sounds (e.g., sighs) and voice qualities like pitch, loudness, speed, and tonal characteristics.

In development, certain milestones mark infant progression along normative developmental pathways. By two months of age, most babies begin to exhibit social responsiveness, such as smiling in reaction to interactions with caregivers (CDC, 2023). By nine months, babies typically demonstrate social engagement by turning their heads and establishing eye contact when called by their name. As approaching their first birthday, they commonly engage in behaviours such as pointing at objects of interest or imitating gestures like waving goodbye (CDC, 2023). However, in cases of autism, non-verbal signals often diverge significantly from typical patterns, with these developmental milestones frequently being delayed or absent. Such deviations are often flagged as early indicators of developmental concerns (National Health Service [NHS], 2022). In fact, the distinct differences in non-verbal signals are so significant that they are one of the essential criteria used in diagnosing autism (APA, 2022; World Health Organization [WHO], 2019). The DSM-5-TR describes these differences in autism as follows (APA, 2022, p. 62–63):

> “Deficits in nonverbal communicative behaviors used for social interaction are manifested by absent, reduced, or atypical use of eye contact (relative to cultural norms), gestures, facial expressions, body orientation, or speech intonation. An early feature of autism spectrum disorder is impaired joint attention as manifested by a lack of pointing, showing, or bringing objects to share interest with others, or failure to follow someone’s pointing or eye gaze. Individuals may learn a few functional gestures, but their repertoire is smaller than that of others, and they often fail to use expressive gestures spontaneously in communication. Among young people and adults with fluent language, the difficulty in coordinating nonverbal communication with speech may give the impression of odd, wooden, or exaggerated “body language” during interactions.”

This thesis explored non-verbal signals at perceptual and production levels, utilising behavioural tasks and measurements. More specifically, predictions about facial expressions (perception) were first investigated to better characterise how autistic adults perceive non-verbal signals. Moreover, the production of non-verbal signals (specifically eye, head, and hand movements) were used to digitally phenotype autism and to test if these could be used as potential diagnostic markers for autism. Therefore, to further contextualise the research conducted, I will now examine relevant non-verbal signals in autism and how these could be utilised in practical applications.

Facial expression perception in autism

Ekman & Friesen (1969) proposed six universal and cross-cultural emotions, as well as developed methods for coding facial expressions (Ekman & Friesen, 1978). According to this perspective, often referred to as the classical view of emotions, specific facial expressions correspond to biologically innate emotions such as sadness shown by a pouting expression and happiness by a smile (for a review, see Barrett et al., 2019).

In autism, a significant amount of research has assumed and followed this view over the last 30 years, investigating how autistic individuals recognise and interpret "universal" facial expressions. These studies typically rely on measures of accuracy, assessing the ability to "correctly" identify stereotypical or prototypical expressions of emotion from a limited set of provided emotion labels (i.e., x-alternative-forced-choice design). A recent systematic review and meta-analysis that included 148 studies and used accuracy as the dependent measure concluded that autistic individuals had a significant overall impairment in recognising all basic emotions (Yeung, 2022). This included difficulties in recognising emotions in faces and in other modalities (such as recognising emotion in voices), impairment in recognising facial attributes with or without an emotion component (e.g., recognising the sex or age), as well as impairments in recognising all individual basic emotion types (Yeung, 2022). The meta-analysis explores how task characteristics play a role, with autistic individuals having more difficulties in recognising complex emotions and showing poorer performance on facial emotion recognition tasks that emphasise holistic (e.g., whole-face stimuli) more than featural (e.g., eyes-only) face processing (Yeung, 2022). In addition, differences in the intensity of facial expression stimuli may influence facial expression recognition ability, with studies suggesting greater difficulties recognising low- but not high-intensity expressions (Keating & Cook, 2021).

While these observations provide valuable insights into the differences in performance between groups of individuals and have shown that autistic individuals tend to underperform on such tasks, recent evidence underscores that emotions are experienced and expressed in highly variable ways. Emotions vary significantly across cultures and situations (Barrett et al., 2019), aligning with constructionist views, such as the Theory of Constructed Emotion (Barrett, 2017), which posits that emotions are not biologically discrete but instead represent categories constructed by applying conceptual knowledge about emotion (learned through one's culture, language, and experience) to make meaning of ongoing sensory information arising from the body (e.g., changes in heart rate or breathing). This theory challenges the classical view that emotions are innate, biologically hardwired responses, suggesting instead that emotions are constructed from more fundamental psychological and physiological components. As a result, this perspective does not emphasise accuracy in emotional expression, but rather highlights that different emotions can share similar face configurations (e.g., smiling could result in a pleasant situation but also in an uncomfortable one) and represent heterogeneous categories of instances, depending on the interplay of individual, cultural, and situational factors (Barrett, 2017).

In this thesis, facial expressions were explored (Publication I), moving away from measuring accuracy and instead assessing the extent to which autistic individuals expected to see stereotypical facial expressions in different emotionally evocative scenarios. This approach aimed to further explore the differences between autistic and control individuals in social perception. Specifically, a social perception behavioural task was employed, where participants were asked to predict facial expressions in given scenarios and then use those predictions for social evaluation, more specifically, likability ratings. Studies on social judgements in autism are scarce and revealed mixed results. Some research found no differences in trustworthiness ratings compared to controls (Caulfield et al., 2014; Ewing et al., 2014; Latimier et al., 2019; Mathersul et al., 2013; Pinkham et al., 2008; Walsh et al., 2016), while other studies found significant differences only when rating naturalistic face stimuli (versus synthetic faces; (Forgeot d'Arc et al., 2016), higher overall ratings of trustworthiness and approachability (Adolphs et al., 2001), or even overrating the trustworthiness of negatively balanced faces in autistic individuals (Couture et al., 2010; Losh et al., 2009). Another study examined the ability of autistic individuals to make a range of social judgments—including trustworthiness, intelligence, attractiveness, approachability, and distinctiveness—from pictures of faces (Philip et al., 2010). Using an accuracy measure with right and wrong answers, the study found that autistic individuals had "deficits" (i.e., less accurate responses) in judging intelligence, attractiveness, approachability, and distinctiveness compared to controls (Philip et al., 2010).

Publication I explored social perception in terms of predictions about facial expressions in autistic and control individuals. Additionally, it was investigated how those predictions influence social evaluation, particularly judgements of likability. For a more extensive rationale, refer to the full publication in section 2.

Eye, head and hand movements

One of the early signs of autism in childhood is atypical production/use of eye contact and body movements, such as the absence, avoidance, and decrease of eye contact, and atypical gestures and motor functions (CDC, 2023; NHS, 2022). In typical development, a 4-week-old infant may start looking at faces and returning a mother's gaze (Malik & Marwaha, 2024), with an increase in face fixations between 5 and 7 weeks (Haith et al., 1977). Similarly, babies typically begin to use communicative gestures like showing, giving, and pointing at objects around 9–10 months of age (Capone & McGregor, 2004). Eye contact and the use of gestures allow people to share experiences (Hamilton, 2016), develop joint attention (Hietanen, 2018), and is critical for language and cognitive development (Farroni et al., 2002; Mundy et al., 2003), as well as the development of social skills and higher levels of social cognition abilities, such as mentalisation, affective evaluation, or empathy (Stephenson et al., 2021).

Eye contact and other gaze behaviours have been notable factors in recognising autistic features and have been extensively investigated. Studies include measurements of gaze duration and fixation/allocation, saccade amplitude, blink rate, pupil diameter and dilation, gaze path, or fixation count (e.g., Bast et al., 2019; Camero et al., 2021; Kim et al., 2024; Klin et al., 2002; Krishnappa Babu et al., 2023; Nayar et al., 2022; Tenenbaum et al., 2021; Tortelli et al., 2022;

Wang et al., 2018; Wen et al., 2022). One of the most popular tools to investigate gaze behaviours is eye-tracking, and studies typically include paradigms of videos or photographs of social (e.g., human faces and biological motion) versus non-social stimuli (e.g., objects). Klin et al. (2002) investigated visual fixation patterns in autism while participants viewed social scenes. By using eye-tracking and defining four areas of interest (eyes, mouths, bodies, and objects), the authors found that autistic individuals had reduced fixation times on the eyes and increased salience of mouths, bodies, and objects. They also found that fixation times on mouths and objects (but not eyes) were strong predictors of social competence. In line with these results, Camero and colleagues (2021) found that autistic children looked less (and for shorter times) at the eye region of a face stimulus compared to control individuals. Conversely, autistic children looked at the mouth region a greater number of times and spent more time looking at images of non-existent invented objects, so-called pseudo-objects. This was partially supported by a systematic review and meta-analysis that found that autistic individuals have significant impairments in gaze fixation to the eyes, while no statistically significant variance was observed regarding fixation on the mouth region (Papagiannopoulou et al., 2014). However, the absence of significant differences in mouth fixation may be attributed to methodological disparities across the studies reviewed, as highlighted by the authors (Papagiannopoulou et al., 2014). Another recent systematic review and meta-analysis across ages supports associations between reduced gaze to the face, head, and eye regions with greater social difficulties and increased autism symptom severity (Riddiford et al., 2022). Overall, gaze allocation to the mouth area appeared dependent on the social and emotional content of the scenes. For instance, Wieckowski & White (2017) found higher fixation times on the mouth region specifically in response to surprise stimuli. Krishnappa Babu et al. (2023) examined the blink rate of children while they engaged with a screen, watching both short social and non-social movies. Their findings revealed that autistic children exhibited diminished screen-facing time, and a higher average blink rate compared to control children. Additionally, autistic children demonstrated decreased screen orientation during social films compared to non-social ones, and their blink rate did not differ significantly between social and non-social movies, in contrast to controls, who exhibited the opposite pattern. Consistently, another study reported that autistic individuals allocated less time to gazing at human faces and exhibited prolonged fixation on non-social features within the background (Kim et al., 2024).

Similarly, studies have revealed significant variations in head movements (Campbell et al., 2018; Dawson et al., 2018; Martin et al., 2018; Perochon et al., 2021; Zhao et al., 2021, 2022), as well as in hand and body gestures and movements among autistic individuals (Ardalan et al., 2019; B. Li et al., 2017; Zhao, 2022). Research indicates that autistic toddlers often display reduced responsiveness to their name being called (i.e., turning their heads less to the person calling) and exhibit slower completion of actions compared (e.g., higher latency before turning their heads) to typically developing peers (Campbell et al., 2018; Perochon et al., 2021). Furthermore, autistic children frequently demonstrate non-communicative repetitive behaviours such as head shaking and increased head movement during face-to-face interactions, compared to a control group of children (Zhao et al., 2021). Martin et al. (2018) observed that autistic children exhibit greater head turning, faster head movements, and more lateral inclinations when exposed to social stimulus videos. Dawson et al. (2018) found that autistic toddlers have a notably higher rate of head movement, suggesting challenges in maintaining head position during tasks requiring focused attention. Caballero et al. (2020) investigated involuntary head movements in autistic individuals during fMRI scans, identifying elevated levels across age groups compared to controls. Recent findings by Zhao et al. (2022) highlight increased body movement and reduced movement complexity among autistic children during face-to-face conversations. Ardalan et al. (2019) studied kinematic and postural sway data during video game play, revealing more erratic movements in autistic compared to control young individuals. Additionally, Li et al. (2017) identified differences in kinematic movements during a hand movement imitation task between autistic and control participants.

As seen, research shows robust differences in eye and body movements among autistic individuals compared to control counterparts. In the present thesis, these known specific non-verbal communication signals were also explored, this time from an applied perspective (Publication II). More specifically, a system was developed to automatically gather the eye, head, and hand movements of autistic and control individuals while they were engaged in a VR task, aiming to determine the feasibility of this system to capture such movements. Then, the potential use of ML techniques for pattern classification was explored. The evidence of these technologies in autism and their potential relevance will now be reviewed.

Applications of non-verbal signals for pattern classification in autism

Non-verbal signals have been explored as potential markers for the diagnosis of autism (for a comprehensive overview of biomarkers in general, see the systematic review by Parellada et al., 2022). This topic remains a current focus in research due to several compelling reasons. To date, the cause of autism has not yet been identified (for a recent review, see Hodges et al., 2020), and therefore, the current diagnosis of autism relies entirely on clinical assessment of diagnostic criteria (National Institute for Health and Care Excellence [NICE], 2013). The diagnosis of autism is a complex and resource-intensive clinical process that often requires a multidisciplinary team, including autism experts. This comprehensive evaluation typically includes anamnestic interviews to gather comprehensive developmental information, a neuropsychological evaluation, and the application of autism-specific standardised tests (for comprehensive guidelines, see the German S3-guidelines, AWMF, 2015, the European Society of Child and Adolescent Psychiatry [ESCAP] guideline, Fuentes et al., 2021, or the NICE guideline, 2012). The autism-specific tests, such as the Autism Diagnostic Interview-Revised (ADI-R, Rutter et al., 2003) and the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2, Lord et al., 2012), have demonstrated good specificity and sensitivity levels in children (Frazier et al., 2024; for a recent systematic review and meta-analysis, see Lebersfeld et al., 2021), but they are less reliable in adolescents and adults (AWMF, 2015). Additionally, a careful differential diagnosis is necessary to ensure an accurate and reliable diagnosis of autism. These complexities underscore the pressing need for advancements in diagnostic approaches.

In the medical field, ML, a branch of artificial intelligence, has gained significant attention for its ability to analyse data and predict patterns/differences across groups. In short, ML is a collection of techniques that focus specifically on algorithms and statistical models that allow computers to learn from data, recognise patterns, and make decisions or predictions (for an introduction to ML, see Bishop, 2006). ML has successfully (i.e., a model makes accurate predictions or classifications while being efficient and scalable) been applied to objectively classify different kinds of diseases and conditions, like skin cancer (Codella et al., 2015; Esteva et al., 2017) or heart diseases (Rajpurkar et al., 2017). There is now increasing attention on applying ML to mental conditions, such as psychosis (Koutsouleris et al., 2021).

In recent years, there has been a growing development of computer-aided investigations of autism through ML. Researchers in autism have explored various domains, such as brain activity (e.g., for a review of fMRI data, see Liu et al., 2021), genetics (for a systematic review, see Bracher-Smith et al., 2021), eye movements (for a recent systematic review and meta-analysis, see Wei et al., 2023), and other non-verbal signals such as body movements (e.g., Ardalan et al., 2019; Li et al., 2017; Zhao et al., 2022), voice features (Ochi et al., 2019; Plank et al., 2023; Santos et al., 2013), interpersonal synchrony (Georgescu et al., 2019; Koehler, Dong, Bierlich, et al., 2024; Koehler, Dong, Song, et al., 2024; Koehler et al., 2021), or facial behaviour (Derbali et al., 2023; Drimalla et al., 2020). The usability of digital applications is actively discussed (Georgescu et al., 2019; Koehler & Falter-Wagner, 2023). In particular, the use of VR technologies in autism research has increased due to several advantages that it offers, such as a strong level of experimental control (Parsons, 2015).

VR offers a computer-generated simulation of a three-dimensional image or environment that can be interacted with in a manner that feels real or physical to the user, facilitated by specialised electronic equipment like head-mounted displays with screens or gloves equipped with sensors (Mandal, 2013). VR systems provide ecologically valid environments that may be engaging, motivating, and enjoyable, while also allowing for the collection of behavioural performance data during gameplay (Alcañiz et al., 2022). Recent studies have highlighted its potential for enhancing social skills, emotional regulation, and other areas critical for autistic individuals. For example, a recent systematic review and meta-analysis by Li et al. (2023) indicates that VR interventions can effectively support social and affective skill training in autistic individuals by providing immersive and controlled environments for practice. Additionally, VR has been applied in interventions for autistic individuals in emotion regulation (Lorenzo et al., 2016; Yuan & Ip, 2018), to treat phobias (Maskey et al., 2014, 2019), or to practice fine motor skills (Zhao et al., 2018), among others. Another review by Savickaite et al. (2022) suggests that VR's unique capabilities, such as immersion and interactivity, make it a valuable tool for autism research. The authors also emphasise that most research is heavily concentrated on intervention work, particularly in social training (Savickaite et al., 2022), and that other important aspects, such as applications of VR in diagnosis, have yet to be thoroughly explored.

Since VR technologies have been tested as diagnostic tools to aid evaluation processes in populations with psychiatric conditions such as psychosis (for a review, see Lan et al., 2023) and other neurodevelopmental disorders such as attention deficit hyperactivity disorder (Areces et al., 2018a, 2018b), VR technologies also represent a potential application to aid the autism diagnosis. This thesis explored the potential applicability of non-verbal communication signals to digitally phenotype autism. Specifically, in Publication II, a VR-based system that automatically extracts eye, hand, and head movements was used for ML and pattern classification. Importantly, the purpose of this study was not to provide an ML classification algorithm to be readily translatable into clinical practice, as for this purpose several steps of investigation including large autism samples with clinical and non-clinical comparison samples for replication and additional validation samples would be required. Instead, the purpose was to create and show the potentially viable applicability of an immersive VR system designed to capture known non-verbal communication differences in a controlled setting with minimal noise, allowing for precise data extraction and potential ML classification application, i.e., a feasibility or proof-of-concept study. For a more extensive rationale, refer to the full publication in section 3.

This introductory section aimed to provide a short overview of the state-of-the art literature on non-verbal signals in autism and to contextualise the exploratory and applicability work included in this thesis. The next section will detail the aims and hypotheses of the two independent studies conducted.

1.1. Aims and hypotheses

The general aim of this thesis was to explore non-verbal signals in autism. Concretely, the research questions (RQ) of this thesis were: RQ1: How are non-verbal signals being perceived and used for social evaluation in autistic adults? and RQ2: Could non-verbal communication signals be objectively measured in a VR setting and be used for pattern classification? To explore these RQ, two independent studies were conducted (Publications I and II, respectively).

Publication I used a social perception behavioural paradigm (Chanes et al., 2018), aiming to explore facial expression predictions in autistic individuals and to assess the extent to which they predict encountering stereotypical facial expressions in emotionally evocative contexts. Additionally, we aimed to study whether these predictions impacted social evaluation in autism. In previous studies using this task, adult participants without any psychiatric condition or medication were shown to hold stereotypical predictions about facial expressions in emotion contexts (Chanes et al., 2018): participants rated stereotypical facial expressions matching the emotion evoked by the scenario (e.g., a smiling face for a happy scenario) as more similar to what they had imagined than when the stereotypical facial expression did not match the emotion evoked by the preceding scenario (e.g., a smiling face for a sad scenario). This task has also been used with clinical populations before, particularly with individuals with depression (Ramos-Grille et al., 2022). Since autistic individuals often present differences (mainly described as reduced abilities) in emotion recognition (Harms et al., 2010; Lozier et al., 2014; Uljarevic & Hamilton, 2013; Yeung, 2022), it was hypothesised (H) that (H1) predictions about facial expressions would be less stereotypical for autistic than for control individuals. Moreover, since autism is characterised by differences (or deficits, from a medical-perspective model) in social interaction (APA, 2022), it was also hypothesised that (H2) such reduced stereotypically in autistic individuals could be associated with differences in social evaluation.

Publication II involved the development of a VR-based system designed for capturing non-verbal communication signals and employing ML algorithms to classify autism behaviours. The primary aim was to explore the potential of eye, head, and hand movements as markers of autism in a virtual simulation to maximise the experimental control and further use these data for pattern classification. Since several studies showed that autistic individuals have reduced eye fixation on the eye area of a face stimulus and instead show increased eye fixation in the mouth area of a face stimulus and in background areas compared to control individuals (Camero et al., 2021; Kim et al., 2024; Klin et al., 2002), it was hypothesised that (H3) there would be a difference in mean fixation times on the eye, mouth, and background areas during social interaction with a virtual character between autistic and control individuals. Moreover, since other non-verbal communication signals also seem to be different in autism (Ardalan et al., 2019; Dawson et al., 2018; Li et al., 2017; Zhao et al., 2021, 2022), and there are other studies showing the potential of VR, ML, and diagnostics (Areces et al., 2018a, 2018b), it was hypothesised that (H4) the classification of autism versus control individuals based on non-verbal signals (gaze, head, and hand movements) acquired through such a system was feasible.

The following section summarises Publications I and II. Full publications can be found in sections 2 and 3, respectively.

1.2. Summaries of the publications

Publication I: Reduced stereotypicality and spared use of facial expression predictions for social evaluation in autism

The original research reported here investigated the role of facial expression predictions for social evaluation in autism. Using a social perception task, the study found that, although autistic individuals exhibited relatively stereotypical predictions about facial expressions (expecting, e.g., a pouting face after a scenario evoking sadness), they did so to a lesser extent than did control individuals. Interestingly, this reduction of stereotypicality did not impact their use of facial expression predictions in social evaluation (judgements of likability). Indeed, both groups relied to the same extent on their predictions about facial expressions to rate the likability of others. Thus, despite differences in predictions, autistic individuals appear to use them similarly as control individuals for social evaluation. Please refer to section 2 for the full publication.

Publication II: A Virtual Reality Based System for the Screening and Classification of Autism

The original work reported here explored the feasibility of an immersive VR system designed to capture known non-verbal communication differences between autistic and control individuals. Participants were presented with a virtual supermarket scenario where participants interacted with a virtual character acting as a shop seller. The VR system recorded their non-verbal communication behaviours during tasks like shopping and interacting with the virtual character. Data acquisition involved tracking gaze, head, and hand movements, which were then analysed for pattern classification using ML techniques. The study showed the feasibility of extracting expected significant differences in gaze fixation times and head movements between groups of individuals. The VR system used enabled precise data extraction in a controlled environment with minimal noise, making it suitable for potential ML classification applications. Please refer to section 3 for the full publication.

1.3. Discussion

Non-verbal signals have been the central focus of the research presented in the current thesis and have been investigated in autistic and control adults at both perceptual and production levels (Publications I and II). The presented research explored non-verbal signals (in terms of predictions about facial expressions) in adult autistic individuals at a perceptual level, and assessed if known differences in non-verbal behavioural signals (eye, head, and hand movements) could be objectively measured through VR, in a controlled and systematic manner, for later ML classification applications.

Non-verbal communication signals are a core characteristic of autism (APA, 2022). The overall findings of the research presented in this thesis contribute to the literature in this field. Whether it was at a perceptual level, where autistic participants were asked to predict facial expressions in a given emotion-evoking scenario (Publication I), or when producing eye, head, and body movements when interacting with a virtual character (Publication II), autistic individuals consistently exhibited distinct performance patterns in tasks involving non-verbal behaviours, in line with current literature (Camero et al., 2021; Kim et al., 2024; Löytömäki et al., 2020; Zhao et al., 2021, 2022). As such, differences in non-verbal processing are part of the symptomatic definition of autism and therefore expected and do not present a novel finding per se. However, the novel contribution of the findings presented in this thesis are (1) the exploration of the level of stereotypicality of autistic individuals when making predictions about facial expressions, rather than an approach based on “accuracy”, and their use for social evaluation, and (2) the development and feasibility assessment of a VR system that could automatically capture and analyse produced non-verbal communication signals, and that such analysis could be potentially used later for pattern classification using ML techniques.

At a perceptual level, Publication I showed that autistic individuals generally rated facial expressions as more different to what they expected when the stereotypical facial expression for the evoked emotion was presented. This finding revealed that autistic individuals found stereotypical expressions (e.g., smiling faces in a normatively happy scenario) to be less aligned with their own predictions for what the target person would look like in that scenario than did control individuals, thus confirming Hypothesis 1 (H1: Predictions about facial expressions would be less stereotypical for autistic than for control individuals). Moreover, despite the differences in their ratings, autistic individuals still made predictions in the same direction as control individuals. Specifically, both groups rated the matched trials as more predictable compared to the non-matched trials, which is consistent with prior studies in control and clinical populations (Chanes et al., 2018; Draganov et al., 2023; Ramos-Grille et al., 2022). Still, the results indicate that autistic individuals perceive non-verbal signals differently from control individuals, revealing a reduction in stereotypicality when asking participants to predict facial expressions. It can be speculated that this lower stereotypicality could be observed as “deficits” in emotion recognition paradigms.

This lack of stereotypicality may be in line with the Weak Central Coherence Theory (Happé et al., 2001), which suggests that autistic individuals focus more on specific details rather than the broader context. Similarly, the Reduced Generalisation Theory (Plaisted, 2001) proposes that autistic individuals experience difficulties in generalising knowledge and experiences. Both theories highlight the fragmented way in which autistic individuals process information, impacting the way they perceive social cues and apply generalised knowledge. In the same vein, this lack of stereotypicality observed in autistic individuals, could also be linked to recent predictive coding accounts of autism, which suggest privileged processing of prediction errors in autistic individuals (for a recent review, see (Cannon et al., 2021). An overreliance on prediction error may lead autistic individuals to incorporate many detailed circumstances and features into their predictions (i.e., internal model), leading to their expectations of facial expressions being more detailed and less generalisable or stereotypical than non-autistic individuals’. How this possibility may differ from predictive processing accounts (e.g., the presence of attenuated priors account postulated by Pellicano & Burr, 2012, or the inflexible weighting of prediction errors and undue adjustment of prior beliefs when the volatility of context changes, observed by the work of Allenmark et al., 2021 and Shi et al., 2024) remains to be explored, since the social perception task used in Publication I did not assess the content of participants' predictions, or how predictions may change if, e.g., volatility of context, is manipulated.

Contrary to H2, the differences observed in facial expression perception between autistic and control individuals were not associated with differences in social evaluation. The results indicated that the use of facial expression predictions for social evaluation (i.e., likability ratings) appeared to be intact among autistic individuals. Perceived likability strongly correlated with perceived predictability of facial expressions in both autistic and control individuals, without significant differences between groups observed. This is consistent with previous work investigating control individuals (Chanes et al., 2018). The literature on social judgements and autism remains limited. While some studies have found no significant differences in social judgements (e.g., trustworthiness, dominance, kindness) between autistic and control individuals (Caulfield et al., 2014; Ewing et al., 2014; Forgeot d'Arc et al., 2016; Latimier et al., 2019; Mathersul et al., 2013; Pinkham et al., 2008; Walsh et al., 2016), other studies do report significant differences in trustworthiness (concretely, an overrating; Adolphs et al., 2001) or in other constructs such as intelligence or attractiveness (Couture et al., 2010; Losh et al., 2009; Philip et al., 2010). Previous studies often rely on accuracy measures with right and wrong answers (e.g. Philip et al., 2010). Philip and colleagues suggest that the deficits seen in emotion recognition could be part of a broader deficit in mental state attribution in autism, since they found differences between control and autistic individuals both in emotion recognition and social judgements using accuracy measures. Interestingly, based on the present results, it can be speculated that considering individuals' own expectations when making social judgments, rather than relying on prespecified assumptions about accuracy, might account for the lack of observed differences between autistic and control individuals. Moreover, factors such as stimulus complexity, individual variability in interpreting social signals, or the context in which judgements are made can all influence outcomes. Future research should incorporate a broader range of social contexts and stimuli, as well as account for individual variability in social perception and evaluation. Additionally, integrating both qualitative and quantitative methods could offer deeper insights into how autistic and control individuals process and evaluate social information. This could lead to more tailored interventions and support strategies that account for individual differences in social processing.

Literature has shown that autistic individuals often present reduced production of non-verbal communication behaviours during social interactions (APA, 2022). Publication II examined eye, head and hand movements in autistic and control individuals in a VR setting, aiming to capture these known differences between autistic and control individuals in a controlled setting with minimal noise, allowing for precise data extraction and potential ML classification application. Partially confirming the proposed hypothesis (H3), results revealed distinct patterns in eye movements between autistic and control participants. Specifically, the study found that autistic individuals exhibited different eye movement patterns when interacting with social stimuli, such as reduced fixation on the eyes of an interaction partner (a VR character) and increased attention to background areas (i.e., non-social stimuli), compared to control participants. These results are in line with clinical descriptions of autism that report reduced eye contact during social interactions (APA, 2022; CDC, 2023; NHS, 2022; WHO, 2019). Reduced fixation time in the eyes region of the virtual character is also supported by recent studies and meta-analyses that found decreased orientation to the eyes region (Camero et al., 2021; Papagiannopoulou et al., 2014, respectively) and by results reported by Kim et al. (2024), who reported longer fixation times on background stimuli. However, contrary to H3 and other studies (Camero et al., 2021; Neumann et al., 2006; Riddiford et al., 2022; Wieckowski & White, 2017), no significant differences across groups were found in the mean fixation times on the mouth area of the virtual character. Still, our results were supported by a systematic review and meta-analysis encompassing six studies that suggested no significant differences in mouth fixation (Papagiannopoulou et al., 2014). The results of Publication II also suggested no significant differences between groups in the mean fixation times in the overall head area of the virtual character. However, autistic individuals exhibited less overt attention towards the virtual character and more toward the background, such as objects in the supermarket, which corresponds with findings of reduced social orientation in previous studies (Kim et al., 2024; Krishnappa Babu et al., 2023).

Regarding data of the participants' head and hand movements, Publication II detected significant differences between groups in only one of the computed features (head rotation average). Although a comprehensive comparison was not conducted, autistic individuals appeared to exhibit greater head movement, consistent with recent research (Martin et al., 2018; Roth et al., 2020; Zhao et al., 2022). Lastly, hand movement data was not analysed in group comparisons due to difficulties in registering the data accurately. Participants were holding two controllers, which restricted natural hand movements, resulting in poor data quality. As a result, this data was deemed unreliable for meaningful group analysis and was excluded from the comparisons.

Publication II confirmed the feasibility of capturing non-verbal communication data within a VR setting and using this data for pattern classification. The effect size revealed that eye movement emerged as a particularly strong potential feature for distinguishing autism, in line with a recent systematic review and meta-analysis (Wei et al., 2023). The VR system proved effective for controlled collection of non-verbal communication signatures, and ML techniques were successfully employed. Although high classification accuracies were obtained, these were not the focus of the current proof-of-concept study as a valid interpretation of high classification accuracies will require future collection of large testing and validation samples. Importantly, the unique system proposed in Publication I—a single-user VR setup designed for social interaction—offered high experimental control. As such the project delivered a platform for future application in marker research allowing exploration of non-verbal communication signals in a controlled and reliable way. Nevertheless, the measured non-verbal signatures contribute to the existing literature that identifies eye movements as a robust marker for autism (for a systematic review, see Parellada et al., 2022) and also in VR settings (Alcañiz et al., 2020, 2022). In addition, including individuals with other conditions, such as borderline personality disorder, could help confirm whether the observed patterns are specific to autism and not attributable to other disorders. Ultimately, this research has the potential to support the development of objective diagnostic tools for more accurate and reliable autism assessments in the future.

The current thesis and the included studies present several limitations that should be addressed in future research. First, while the studies investigated non-verbal signals, they did not explore all possible factors influencing these signals, such as the specifics of expectations or other contextual variables. In Publication I, participants were only asked whether their predictions were similar to the actual displayed expressions, without assessment of the content of the predictions themselves. Future work may characterise autistic individuals' predictions with regards to controls and assess whether differences exist with regards to clarity, level of detail, or other specific aspects. Moreover, the task used in Publication I only assessed three emotion categories (happiness, sadness, and fear) and one social evaluation aspect (likability). Other emotions may be assessed in future studies to better explore and characterise the observed effects. Similarly, other ratings may be used in the future to explore social evaluation more thoroughly (e.g., trustworthiness, Chanes et al., 2018). Furthermore, the aim of Publication I was to assess predictability holistically. Thus, the contribution of different aspects of facial expressions or emotion scenarios to the present findings (e.g., facial configuration, intensity of expression, specific cognitive aspects of image perception, narrative interpretations, affective response to narratives, etc.) remain to be explored in future investigations.

Similarly, Publication II focused only on mean fixation times, whereas other features related to gaze, such as fixation count, saccades lengths or durations, or scan paths were not explored. These may be explored in future research to provide a more comprehensive characterisation of gaze patterns in autism. Additionally, in Publication II, participants were required to use two controllers and wear a head-mounted display, which limited their freedom of hand and head movements. Future research could benefit from incorporating more advanced devices that offer controller-free hand tracking and utilise lighter, less restrictive head-mounted displays. Such advancements could facilitate capturing other known differences in non-verbal communication signals in autism, such as reduced use of head and hand communicative gestures (APA, 2022).

Finally, participant diversity is related to the generalisability of the findings. Future studies may consider more diverse samples to enhance the applicability of the results to a wider population, including individuals with a broader range of intellectual capabilities, co-occurrences, and varying ages across the lifespan.

For a more detailed discussion of the independent studies, please see the full study publications in Sections 2 and 3.

1.4. Conclusions

This thesis presents a comprehensive investigation into the non-verbal signals of autistic individuals, focusing on both perceptual and production aspects.

The findings revealed differences between autistic and control adults, particularly in the way facial expressions are anticipated. Specifically, autistic individuals tended to experience facial expressions as less predictable compared to controls, indicating a deviation from typical expectations about emotion facial configuration stereotypes. Despite this, the use of facial expression predictions for social evaluation (judgements of likability) appeared to be intact among autistic individuals.

Furthermore, this work showed the development and feasibility assessment of a single-user VR system designed to capture non-verbal communication signals. The VR setup provided high experimental control, facilitating the systematic evaluation of non-verbal cues in a controlled environment. This system not only proved to be an effective tool for capturing known non-verbal communication signals of autism, but also provided a platform for future applications in marker research and ML classification.

Overall, this thesis contributes to a more nuanced understanding of non-verbal signals in autism and highlights the potential for innovative technological tools to enhance research and application in this field.

1.5. References

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. In *Journal of Cognitive Neuroscience* (Vol. 13, Issue 2, pp. 232–240). MIT Press. https://doi.org/10.1162/089892901564289

Alcañiz, M., Chicchi Giglioli, I. A., Sirera, M., Minissi, E., & Abad, L. (2020). [Autism spectrum disorder biomarkers based on biosignals, virtual reality and artificial intelligence]. *Medicina*, *80 Suppl 2*, 31–36. https://www.medicinabuenosaires.com/PMID/32150710.pdf

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

Allenmark, F., Shi, Z., Pistorius, R. L., Theisinger, L. A., Koutsouleris, N., Falkai, P., Müller, H. J., & Falter-Wagner, C. M. (2021). Acquisition and Use of 'Priors' in Autism: Typical in Deciding Where to Look, Atypical in Deciding What Is There. *Journal of Autism and Developmental Disorders*, *51*(10), 3744–3758. https://doi.org/10.1007/s10803-020-04828-2

American Psychiatric Association (APA). (2022). *Diagnostic and statistical manual of mental disorders (5th ed., text rev.)*. https://doi.org/10.1176/appi.books.9780890425787

Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). (2015). *S3-Leitlinie Autismus-Spektrum-Störungen im Kindes-, Jugend- und Erwachsenenalter Teil 1: Diagnostik*. DGKJP, DGPPN. https://www.awmf.org/leitlinien/detail/ll/028-018.html

Ardalan, A., Assadi, A. H., Surgent, O. J., & Travers, B. G. (2019). Whole-Body Movement during Videogame Play Distinguishes Youth with Autism from Youth with Typical Development. *Scientific Reports*, *9*(1), 20094. https://doi.org/10.1038/s41598-019-56362-6

Areces, D., Dockrell, J., García, T., González-Castro, P., & Rodríguez, C. (2018). Analysis of cognitive and attentional profiles in children with and without ADHD using an innovative virtual reality tool. *PloS One*, *13*(8), e0201039. https://doi.org/10.1371/journal.pone.0201039

Areces, D., Rodríguez, C., García, T., Cueli, M., & González-Castro, P. (2018). Efficacy of a continuous performance test based on virtual reality in the diagnosis of ADHD and its clinical presentations. *Journal of Attention Disorders*, *22*(11), 1081–1091. https://doi.org/10.1177/1087054716629711

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Bast, N., Banaschewski, T., Dziobek, I., Brandeis, D., Poustka, L., & Freitag, C. M. (2019). Pupil dilation progression modulates aberrant social cognition in autism spectrum disorder. *Autism Research*, *12*(11), 1680–1692. https://doi.org/10.1002/aur.2178

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer New York, NY.

Bosman, R., & Thijs, J. (2024). Language Preferences in the Dutch Autism Community: A Social Psychological Approach. *Journal of Autism and Developmental Disorders*, *54*(5), 1727–1739. https://doi.org/10.1007/s10803-023-05903-0

Botha, M., Dibb, B., & Frost, D. M. (2022). "Autism is me": an investigation of how autistic individuals make sense of autism and stigma. *Disability & Society*, *37*(3), 427–453. https://doi.org/10.1080/09687599.2020.1822782

Bracher-Smith, M., Crawford, K., & Escott-Price, V. (2021). Machine learning for genetic prediction of psychiatric disorders: a systematic review. *Molecular Psychiatry*, *26*(1), 70–79. https://doi.org/10.1038/s41380-020-0825-2

Buijsman, R., Begeer, S., & Scheeren, A. M. (2023). 'Autistic person' or 'person with autism'? Person-first language preference in Dutch adults with autism and parents. *Autism*, *27*(3), 788–795. https://doi.org/10.1177/13623613221117914

Bury, S. M., Jellett, R., Haschek, A., Wenzel, M., Hedley, D., & Spoor, J. R. (2023). Understanding language preference: Autism knowledge, experience of stigma and autism identity. *Autism*, *27*(6), 1588–1600.

Bury, S. M., Jellett, R., Spoor, J. R., & Hedley, D. (2020). "It defines who I am" or "It's something I have": What language do [autistic] Australian adults [on the autism spectrum] prefer? *Journal of Autism and Developmental Disorders*, *53*(2), 677–687. https://doi.org/10.1007/s10803-020-04425-3

Caballero, C., Mistry, S., & Torres, E. B. (2020). Age-Dependent Statistical Changes of Involuntary Head Motion Signatures Across Autism and Controls of the ABIDE Repository. *Frontiers in Integrative Neuroscience*, *14*, 23. https://doi.org/10.3389/fnint.2020.00023

Camero, R., Martínez, V., & Gallego, C. (2021). Gaze Following and Pupil Dilation as Early Diagnostic Markers of Autism in Toddlers. *Children*, *8*(2), 113. https://doi.org/10.3390/children8020113

Campbell, K., Carpenter, K. L. H., Hashemi, J., Espinosa, S., Marsan, S., Borg, J. S., Chang, Z., Qiu, Q., Vermeer, S., Adler, E., Tepper, M., Egger, H. L., Baker, J. P., Sapiro, G., & Dawson, G. (2018). Computer vision analysis captures atypical attention in toddlers with autism. *Autism*, *23*(3), 619–628. https://doi.org/10.1177/1362361318766247

Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. *Autism Research*, *14*(4), 604–630. https://doi.org/10.1002/aur.2482

Capone, N., & McGregor, K. (2004). Gesture Development. *Journal of Speech, Language, and Hearing Research*, *47*(1), 173–186. https://doi.org/10.1044/1092-4388(2004/015)

Caulfield, F., Ewing, L., Burton, N., Avard, E., & Rhodes, G. (2014). Facial Trustworthiness Judgments in Children with ASD Are Modulated by Happy and Angry Emotional Cues. *PLOS ONE*, *9*(5), e97644-. https://doi.org/10.1371/journal.pone.0097644

Centers for Disease Control and Prevention. (2023). *CDC's Developmental Milestones*. https://www.cdc.gov/ncbddd/actearly/milestones/index.html

Chanes, L., Wormwood, J. B., Betz, N., & Barrett, L. F. (2018). Facial expression predictions as drivers of social perception. *Journal of Personality and Social Psychology*, *114*(3), 380. https://doi.org/10.1037/pspa0000108

Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., & Smith, J. R. (2015). *Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images* (L. Zhou, L. Wang, Q. Wang, & Y. Shi, Eds.; pp. 118–126). Springer International Publishing. https://doi.org/10.1017/S003329170999078X

Couture, S. M., Penn, D. L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: more convergence than divergence. *Psychological Medicine*, *40*(4), 569–579. https://doi.org/10.1017/S003329170999078X

Dawson, G., Campbell, K., Hashemi, J., Lippmann, S. J., Smith, V., Carpenter, K., Egger, H., Espinosa, S., Vermeer, S., Baker, J., & Sapiro, G. (2018). Atypical postural control can be detected via computer vision analysis in toddlers with autism spectrum disorder. *Scientific Reports*, *8*(1), 17008. https://doi.org/10.1038/s41598-018-35215-8

De Laet, H., Nijhof, A. D., & Wiersema, J. R. (2023). Adults with Autism Prefer Person-First Language in Dutch: A Cross-Country Study. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-023-06192-3

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., Hickie, I., & Guastella, A. J. (2018). Autism spectrum disorders: a meta-analysis of executive function. *Molecular Psychiatry*, *23*(5), 1198–1204. https://doi.org/10.1038/mp.2017.75

Derbali, M., Jarrah, M., & Randhawa, P. (2023). Autism spectrum disorder detection: Video games based facial expression diagnosis using deep learning. *International Journal of Advanced Computer Science and Applications*, *14*(1). https://doi.org/10.14569/IJACSA.2023.0140112

Di Vara, S., Guerrera, S., Menghini, D., Scibelli, F., Lupi, E., Valeri, G., & Vicari, S. (2024). Characterizing individual differences in children and adolescents with autism spectrum disorder: a descriptive study. *Frontiers in Psychology*, *15*, 1323787. https://doi.org/10.3389/fpsyg.2024.1323787

Draganov, M., Galiano-Landeira, J., Doruk Camsari, D., Ramírez, J.-E., Robles, M., & Chanes, L. (2023). Noninvasive modulation of predictive coding in humans: causal evidence for frequency-specific temporal dynamics. *Cerebral Cortex*, *33*(13), 8421–8430. https://doi.org/10.1093/cercor/bhad127

Drimalla, H., Scheffer, T., Landwehr, N., Baskow, I., Roepke, S., Behnia, B., & Dziobek, I. (2020). Towards the automatic detection of social biomarkers in autism spectrum disorder: introducing the simulated interaction task (SIT). *Npj Digital Medicine*, *3*(1), 1–10. https://doi.org/10.1038/s41746-020-0227-5

Ekman, P., & Friesen, W. V. (1969). *The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding*. *1*(1), 49–98. https://doi.org/10.1515/semi.1969.1.1.49

Ekman, P., & Friesen, W. V. (1978). Facial action coding system. *Environmental Psychology & Nonverbal Behavior*. https://doi.org/10.1037/t27734-000

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2014). Appearance-based trust behaviour is reduced in children with autism spectrum disorder. *Autism*, *19*(8), 1002–1009. https://doi.org/10.1177/1362361314559431

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Forgeot d'Arc, B., Ramus, F., Lefebvre, A., Brottier, D., Zalla, T., Moukawane, S., Amsellem, F., Letellier, L., Peyre, H., Mouren, M.-C., Leboyer, M., & Delorme, R. (2016). Atypical Social Judgment and Sensitivity to Perceptual Cues in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *46*(5), 1574–1581. https://doi.org/10.1007/s10803-014-2208-5

Frazier, T. W., Whitehouse, A. J. O., Leekam, S. R., Carrington, S. J., Alvares, G. A., Evans, D. W., Hardan, A. Y., & Uljarević, M. (2024). Reliability of the Commonly Used and Newly-Developed Autism Measures. *Journal of Autism and Developmental Disorders*, *54*(6), 2158–2169. https://doi.org/10.1007/s10803-023-05967-y

Friedman, A., & Sterling, L. (2019). A Review of Language, Executive Function, and Intervention in Autism Spectrum Disorder. *Seminars in Speech and Language*, *40*(04), 291–304. https://doi.org/10.1055/s-0039-1692964

Fuentes, J., Hervás, A., & Howlin, P. (2021). ESCAP practice guidance for autism: a summary of evidence-based recommendations for diagnosis and treatment. *European Child & Adolescent Psychiatry*, *30*(6), 961–984. https://doi.org/10.1007/s00787-020-01587-4

Geelhand, P., Papastamou, F., Belenger, M., Clin, E., Hickman, L., Keating, C. T., & Sowden, S. (2023). Autism-Related Language Preferences of French-Speaking Autistic Adults: An Online Survey. *Autism in Adulthood*, *5*(3), 275–288. https://doi.org/10.1089/aut.2022.0056

Georgescu, A. L., Koehler, J. C., Weiske, J., Vogeley, K., Koutsouleris, N., & Falter-Wagner, C. (2019). Machine Learning to Study Social Interaction Difficulties in ASD. *Frontiers in Robotics and AI*, *0*, 132. https://doi.org/10.3389/frobt.2019.00132

Ghanouni, P., & Quirke, S. (2023). Resilience and Coping Strategies in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *53*(1), 456–467. https://doi.org/10.1007/s10803-022-05436-y

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye Contact and Face Scanning in Early Infancy. *Science*, *198*(4319), 853–855. https://doi.org/10.1126/science.918670

Hall, J. A., & Knapp, M. L. (2013). *Nonverbal communication edited by Judith A. Hall and Mark L. Knapp.* De Gruyter Mouton. https://doi.org/10.1515/9783110238150

Hamilton, A. F. de C. (2016). Gazing at me: the importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150080. https://doi.org/10.1098/rstb.2015.0080

Happé, F., & Baron-Cohen, S. (2014, July 15). *Remembering Lorna Wing (1928-2014)*. https://www.thetransmitter.org/spectrum/remembering-lorna-wing-1928-2014/

Happé, F., Briskman, J., & Frith, U. (2001). Exploring the Cognitive Phenotype of Autism: Weak "Central Coherence" in Parents and Siblings of Children with Autism: I. Experimental Tests. *Journal of Child Psychology and Psychiatry*, *42*(3), 299–307. https://doi.org/10.1111/1469-7610.00723

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

Hietanen, J. K. (2018). Affective Eye Contact: An Integrative Review. *Frontiers in Psychology*, *9*, 1587. https://doi.org/10.3389/fpsyg.2018.01587

Hodges, H., Fealko, C., & Soares, N. (2020). Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. *Translational Pediatrics*, *9*(Suppl 1), S55–S65. https://doi.org/10.21037/tp.2019.09.09

Hull, L., Petrides, K. V, & Mandy, W. (2020). The Female Autism Phenotype and Camouflaging: a Narrative Review. *Review Journal of Autism and Developmental Disorders*, *7*(4), 306–317. https://doi.org/10.1007/s40489-020-00197-9

Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or both? Autism and neurodiversity. *Developmental Psychology*, *49*(1), 59–71. https://doi.org/https://doi.org/10.1037/a0028353

Keating, C. T., & Cook, J. L. (2021). Facial Expression Production and Recognition in Autism Spectrum Disorders: A Shifting Landscape. *Psychiatric Clinics of North America*, *44*(1), 125–139. https://doi.org/10.1016/j.psc.2020.11.010

Keating, C. T., Hickman, L., Leung, J., Monk, R., Montgomery, A., Heath, H., & Sowden, S. (2023). Autism-related language preferences of English-speaking individuals across the globe: A mixed methods investigation. *Autism Research*, *16*(2), 406–428. https://doi.org/https://doi.org/10.1002/aur.2864

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2015). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Kim, N. Y., He, J., Wu, Q., Dai, N., Kohlhoff, K., Turner, J., Paul, L. K., Kennedy, D. P., Adolphs, R., & Navalpakkam, V. (2024). Smartphone-based gaze estimation for in-home autism research. *Autism Research*, *17*(6), 1140–1148. https://doi.org/10.1002/aur.3140

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual Fixation Patterns During Viewing of Naturalistic Social Situations as Predictors of Social Competence in Individuals With Autism. *Archives of General Psychiatry*, *59*(9), 809–816. https://doi.org/10.1001/archpsyc.59.9.809

Koehler, J. C., Dong, M. Sen, Bierlich, A. M., Fischer, S., Späth, J., Plank, I. S., Koutsouleris, N., & Falter-Wagner, C. M. (2024). Machine learning classification of autism spectrum disorder based on reciprocity in naturalistic social interactions. *Translational Psychiatry*, *14*(1), 76. https://doi.org/10.1038/s41398-024-02802-5

Koehler, J. C., Dong, M. Sen, Song, D.-Y., Bong, G., Koutsouleris, N., Yoo, H., & Falter-Wagner, C. M. (2024). Classifying autism in a clinical population based on motion synchrony: a proof-of-concept study using real-life diagnostic interviews. *Scientific Reports*, *14*(1), 5663. https://doi.org/10.1038/s41598-024-56098-y

Koehler, J. C., & Falter-Wagner, C. M. (2023). Digitally assisted diagnostics of autism spectrum disorder. *Frontiers in Psychiatry*, *14*. https://doi.org/10.1038/s41598-024-56098-y

Koehler, J. C., Georgescu, A. L., Weiske, J., Spangemacher, M., Burghof, L., Falkai, P., Koutsouleris, N., Tschacher, W., Vogeley, K., & Falter-Wagner, C. M. (2021). Brief Report: Specificity of Interpersonal Synchrony Deficits to Autism Spectrum Disorder and Its Potential for Digitally Assisted Diagnostics. *Journal of Autism and Developmental Disorders*, *52*, 3718–3726. https://doi.org/10.1007/s10803-021-05194-3

Koutsouleris, N., Dwyer, D. B., Degenhardt, F., Maj, C., Urquijo-Castro, M. F., Sanfelici, R., Popovic, D., Oeztuerk, O., Haas, S. S., Weiske, J., Ruef, A., Kambeitz-Ilankovic, L., Antonucci, L. A., Neufang, S., Schmidt-Kraepelin, C., Ruhrmann, S., Penzel, N., Kambeitz, J., Haidl, T. K., … Meisenzahl, E. (2021). Multimodal Machine Learning Workflows for Prediction of Psychosis in Patients With Clinical High-Risk Syndromes and Recent-Onset Depression. *JAMA Psychiatry*, *78*(2), 195–209. https://doi.org/10.1001/jamapsychiatry.2020.3604

Krishnappa Babu, P. R., Aikat, V., Di Martino, J. M., Chang, Z., Perochon, S., Espinosa, S., Aiello, R., L. H. Carpenter, K., Compton, S., Davis, N., Eichner, B., Flowers, J., Franz, L., Dawson, G., & Sapiro, G. (2023). Blink rate and facial orientation reveal distinctive patterns of attentional engagement in autistic toddlers: a digital phenotyping approach. *Scientific Reports*, *13*(1), 7158. https://doi.org/10.1038/s41598-023-34293-7

Lai, M.-C., & Baron-Cohen, S. (2015). Identifying the lost generation of adults with autism spectrum conditions. *The Lancet Psychiatry*, *2*(11), 1013–1027. https://doi.org/10.1016/S2215-0366(15)00277-1

Lai, M.-C., Lombardo, M. V, & Baron-Cohen, S. (2014). Autism. *The Lancet*, *383*(9920), 896–910. https://doi.org/10.1016/S0140-6736(13)61539-1

Lai, M.-C., & Szatmari, P. (2020). Sex and gender impacts on the behavioural presentation and recognition of autism. *Current Opinion in Psychiatry*, *33*(2), 117–123. https://doi.org/10.1097/YCO.0000000000000575

Lan, L., Sikov, J., Lejeune, J., Ji, C., Brown, H., Bullock, K., & Spencer, A. E. (2023). A Systematic Review of using Virtual and Augmented Reality for the Diagnosis and Treatment of Psychotic Disorders. *Current Treatment Options in Psychiatry*, *10*, 87–107. https://doi.org/10.1007/s40501-023-00287-5

Latimier, A., Kovarski, K., Peyre, H., Fernandez, L. G., Gras, D., Leboyer, M., & Zalla, T. (2019). Trustworthiness and Dominance Personality Traits' Judgments in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(11), 4535–4546. https://doi.org/10.1007/s10803-019-04163-1

Lebersfeld, J. B., Swanson, M., Clesi, C. D., & O'Kelley, S. E. (2021). Systematic Review and Meta-Analysis of the Clinical Utility of the ADOS-2 and the ADI-R in Diagnosing Autism Spectrum Disorders in Children. *Journal of Autism and Developmental Disorders*, *51*(11), 4101–4114. https://doi.org/10.1007/s10803-020-04839-z

Lei, J., Jones, L., & Brosnan, M. (2021). Exploring an e-learning community's response to the language and terminology use in autism from two massive open online courses on autism education and technology use. *Autism*, *25*(5), 1349–1367. https://doi.org/10.1177/1362361320987963

Li, B., Sharma, A., Meng, J., Purushwalkam, S., & Gowen, E. (2017). Applying machine learning to identify autistic adults using imitation: An exploratory study. *PLOS ONE*, *12*(8), e0182652-. https://doi.org/10.1371/journal.pone.0182652

Li, C., Belter, M., Liu, J., & Lukosch, H. (2023). Immersive virtual reality enabled interventions for autism spectrum disorder: A systematic review and meta-analysis. *Electronics*, *12*(11), 2497. https://doi.org/10.3390/electronics12112497

Lord, C., Luyster, R., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule. 2nd. (ADOS-2) Manual (Part II): Toddler Module.* Western Psychological Services.

Lorenzo, G., Lledo, A., Pomares, J., & Roig, R. (2016). Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders. *COMPUTERS & EDUCATION*, *98*, 192–205. https://doi.org/10.1016/j.compedu.2016.03.018

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). Neuropsychological Profile of Autism and the Broad Autism Phenotype. *Archives of General Psychiatry*, *66*(5), 518–526. https://doi.org/10.1001/archgenpsychiatry.2009.34

Löytömäki, J., Ohtonen, P., Laakso, M., & Huttunen, K. (2020a). The role of linguistic and cognitive factors in emotion recognition difficulties in children with ASD, ADHD or DLD. *International Journal of Language & Communication Disorders*, *55*(2), 231–242. https://doi.org/10.1111/1460-6984.12514

Löytömäki, J., Ohtonen, P., Laakso, M., & Huttunen, K. (2020b). The role of linguistic and cognitive factors in emotion recognition difficulties in children with ASD, ADHD or DLD. *International Journal of Language & Communication Disorders*, *55*(2), 231–242.

Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology*, *26*(4), 933–945. https://doi.org/10.1017/S0954579414000479

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V, Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., … Shaw, K. A. (2023). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C. : 2002)*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

Malik, F., & Marwaha, R. (2024). Developmental Stages of Social Emotional Development in Children. In *StatPearls*. http://www.ncbi.nlm.nih.gov/pubmed/27548291

Mandal, F. B. (2014). Nonverbal Communication in Humans. *Journal of Human Behavior in the Social Environment*, *24*(4), 417–421. https://doi.org/10.1080/10911359.2013.831288

Mandal, S. (2013). Brief introduction of virtual reality & its challenges. *International Journal of Scientific & Engineering Research*, *4*(4), 304–309. https://www.ijser.org/researchpaper/Brief-Introduction-of-Virtual-Reality-its-Challenges.pdf

Manusov, V., & Patterson, M. L. (2006). *The Sage handbook of nonverbal communication*. SAGE Publications, Inc. https://doi.org/10.4135/9781412976152

Martin, K. B., Hammal, Z., Ren, G., Cohn, J. F., Cassell, J., Ogihara, M., Britton, J. C., Gutierrez, A., & Messinger, D. S. (2018). Objective measurement of head movement differences in children with and without autism spectrum disorder. *Molecular Autism*, *9*(1), 14. https://doi.org/10.1186/s13229-018-0198-4

Maskey, M., Lowry, J., Rodgers, J., McConachie, H., & Parr, J. R. (2014). Reducing Specific Phobia/Fear in Young People with Autism Spectrum Disorders (ASDs) through a Virtual Reality Environment Intervention. *PLOS ONE*, *9*(7), e100374-. https://doi.org/10.1371/journal.pone.0100374

Maskey, M., McConachie, H., Rodgers, J., Grahame, V., Maxwell, J., Tavernor, L., & Parr, J. R. (2019). An intervention for fears and phobias in young people with autism spectrum disorders using flat screen computer-delivered virtual reality and cognitive behaviour therapy. *Research in Autism Spectrum Disorders*, *59*, 58–67. https://doi.org/10.1016/j.rasd.2018.11.005

Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Psychophysiological correlates of social judgement in high-functioning adults with autism spectrum disorder. *International Journal of Psychophysiology*, *87*(1), 88–94. https://doi.org/10.1016/j.ijpsycho.2012.11.005

Milton, D. (2016). *Re-thinking autism: diagnosis, identity and equality* (K. Runswick-Cole, R. Mallet, & S. Timimi, Eds.; Vol. 31). Jessica Kingsley. https://doi.org/10.1080/09687599.2016.1221666

Monk, R., Whitehouse, A. J. O., & Waddington, H. (2022). The use of language in autism research. *Trends in Neurosciences*, *45*(11), 791–793. https://doi.org/10.1016/j.tins.2022.08.009

Mundy, P., Fox, N., & Card, J. (2003). EEG coherence, joint attention and language development in the second year. *Developmental Science*, *6*(1), 48–54. https://doi.org/10.1111/1467-7687.00253

National Health Service. (2022). *Signs of autism in children*. https://www.nhs.uk/conditions/autism/signs/children/

National Institute for Health and Care Excellence. (2012). *Autism spectrum disorder in adults: diagnosis and management Clinical guideline*. www.nice.org.uk/guidance/cg142

National Institute for Health and Care Excellence. (2013). *Autism spectrum disorder in under 19s: recognition, referral and diagnosis*. https://www.nice.org.uk/Guidance/CG128

National Institute of Mental Health. (2024). *Autism Spectrum Disorder*. https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd

Nayar, K., Shic, F., Winston, M., & Losh, M. (2022). A constellation of eye-tracking measures reveals social attention differences in ASD and the broad autism phenotype. *Molecular Autism*, *13*(1), 18. https://doi.org/10.1186/s13229-022-00490-w

Neumann, D., Spezio, M. L., Piven, J., & Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, *1*(3), 194–202. https://doi.org/10.1093/scan/nsl030

Ochi, K., Ono, N., Owada, K., Kojima, M., Kuroda, M., Sagayama, S., & Yamasue, H. (2019). Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PloS One*, *14*(12), e0225377. https://doi.org/10.1371/journal.pone.0225377

Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience*, *9*(6), 610–632. https://doi.org/10.1080/17470919.2014.934966

Parellada, M., Andreu-Bernabeu, Á., Burdeus, M., San José Cáceres, A., Urbiola, E., Carpenter, L. L., Kraguljac, N. V, McDonald, W. M., Nemeroff, C. B., Rodriguez, C. I., Widge, A. S., State, M. W., & Sanders, S. J. (2022). In Search of Biomarkers to Guide Interventions in Autism Spectrum Disorder: A Systematic Review. *American Journal of Psychiatry*, *180*(1), 23–40. https://doi.org/10.1176/appi.ajp.21100992

Parsons, T. D. (2015). Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00660

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*(10), 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Pellicano, E., & den Houting, J. (2022). Annual Research Review: Shifting from 'normal science' to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry*, *63*(4), 381–396. https://doi.org/10.1111/jcpp.13534

Perochon, S., Di Martino, M., Aiello, R., Baker, J., Carpenter, K., Chang, Z., Compton, S., Davis, N., Eichner, B., Espinosa, S., Flowers, J., Franz, L., Gagliano, M., Harris, A., Howard, J., Kollins, S. H., Perrin, E. M., Raj, P., Spanos, M., … Dawson, G. (2021). A scalable computational approach to assessing response to name in toddlers with autism. *Journal of Child Psychology and Psychiatry*, *62*(9), 1120–1131. https://doi.org/10.1111/jcpp.13381

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., Atkinson, A. P., Calder, A. J., Johnstone, E. C., Lawrie, S. M., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine*, *40*(11), 1919–1929. https://doi.org/10.1017/S0033291709992364

Piaget, J. (1960). The child's conception of the world. In *The child's conception of the world.* Littlefield, Adams.

Pinkham, A. E., Hopfinger, J. B., Pelphrey, K. A., Piven, J., & Penn, D. L. (2008). Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophrenia Research*, *99*(1), 164–175. https://doi.org/10.1016/j.schres.2007.10.024

Plaisted, K. C. (2001). Reduced generalization in autism: An alternative to weak central coherence. In *The development of autism: Perspectives from theory and research.* (pp. 149–169). Lawrence Erlbaum Associates Publishers.

Plank, I. S., Koehler, J. C., Nelson, A., Koutsouleris, N., & Falter-Wagner, C. M. (2023). Automated extraction of speech and turn-taking parameters in autism allows for diagnostic classification using a multivariable prediction model. *Front. Psychiatry*, *14*, 1257569. https://doi.org/10.3389/fpsyt.2023.1257569

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *ArXiv:1711.05225 [Cs, Stat]*. http://arxiv.org/abs/1711.05225

Ramos-Grille, I., Weyant, J., Wormwood, J. B., Robles, M., Vallès, V., Camprodon, J. A., & Chanes, L. (2022). Predictive processing in depression: Increased prediction error following negative valence contexts and influence of recent mood-congruent yet irrelevant experiences. *Journal of Affective Disorders*, *311*, 8–16. https://doi.org/10.1016/j.jad.2022.05.030

Riddiford, J. A., Enticott, P. G., Lavale, A., & Gurvich, C. (2022). Gaze and social functioning associations in autism spectrum disorder: A systematic review and meta-analysis. *Autism Research*, *15*(8), 1380–1446. https://doi.org/10.1002/aur.2729

Ritvo, E. R., & Freeman, B. J. (1984). A medical model of autism: etiology, pathology and treatment. *Pediatric Annals*, *13*(4), 298–305. http://www.ncbi.nlm.nih.gov/pubmed/6610163

Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience*, *18*(11), 671–684. https://doi.org/10.1038/nrn.2017.112

Roth, D., Jording, M., Schmee, T., Kullmann, P., Navab, N., & Vogeley, K. (2020). Towards Computer Aided Diagnosis of Autism Spectrum Disorder Using Virtual Environments. *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 115–122. https://doi.org/10.1109/AIVR50618.2020.00029

Rutter, M., LeCouteur, A., & Lord, C. (2003). *ADI-R. Autism Diagnostic Interview Revised. Manual*. Western Psychological Services.

Santos, J. F., Brosh, N., Falk, T. H., Zwaigenbaum, L., Bryson, S. E., Roberts, W., Smith, I. M., Szatmari, P., & Brian, J. A. (2013). Very early detection of Autism Spectrum Disorders based on acoustic analysis of pre-verbal vocalizations of 18-month old toddlers. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 7567–7571. https://doi.org/10.1109/ICASSP.2013.6639134

Saure, E., Castrén, M., Mikkola, K., & Salmi, J. (2023). Intellectual disabilities moderate sex/gender differences in autism spectrum disorder: A systematic review and meta-analysis. *Journal of Intellectual Disability Research*, *67*(1), 1–34. https://doi.org/10.1111/jir.12989

Savickaite, S., McDonnell, N., & Simmons, D. (2022). *Defining virtual reality (VR). Scoping literature review on VR applications in autism research*. https://doi.org/https://osf.io/p3nh6/

Shakes, P., & Cashin, A. (2019). Identifying Language for People on the Autism Spectrum: A Scoping Review. *Issues in Mental Health Nursing*, *40*(4), 317–325. https://doi.org/10.1080/01612840.2018.1522400

Shi, Z., Allenmark, F., Theisinger, L. A., Pistorius, R. L., Glasauer, S., Müller, H. J., & Falter-Wagner, C. M. (2024). Beyond Prior Belief and Volatility: The Distinct Iterative Prior Updating Process in ASD. *BioRxiv*, 2022.01.21.477218. https://doi.org/10.1101/2022.01.21.477218

Stephenson, L. J., Edwards, S. G., & Bayliss, A. P. (2021). From Gaze Perception to Social Cognition: The Shared-Attention System. *Perspectives on Psychological Science*, *16*(3), 553–576. https://doi.org/10.1177/1745691620953773

Sukhareva, G. E. (1925). Shizoidnyye psixopatii v detskom vozraste [Schizoid personality disorders of childhood]. In M. O. Gurevich (Ed.), *Voprosy pedologii i detskoĭ psikhonevrologii* (2nd ed., pp. 157–187). Zhizn' i Znanie.

Taboas, A., Doepke, K., & Zimmerman, C. (2022). Preferences for identity-first versus person-first language in a US sample of autism stakeholders. *Autism*, *27*(2), 565–570. https://doi.org/10.1177/13623613221130845

Tenenbaum, E. J., Major, S., Carpenter, K. L. H., Howard, J., Murias, M., & Dawson, G. (2021). Distance from Typical Scan Path When Viewing Complex Stimuli in Children with Autism

Spectrum Disorder and its Association with Behavior. *Journal of Autism and Developmental Disorders*, *51*(10), 3492–3505. https://doi.org/10.1007/s10803-020-04812-w

Tepest, R. (2021). The Meaning of Diagnosis for Different Designations in Talking About Autism. *Journal of Autism and Developmental Disorders*, *51*(2), 760–761. https://doi.org/10.1007/s10803-020-04584-3

The World Medical Association. (2013). Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Tortelli, C., Pomè, A., Turi, M., Igliozzi, R., Burr, D. C., & Binda, P. (2022). Contextual information modulates pupil size in autistic children. *Frontiers in Neuroscience*, *16*, 752871. https://doi.org/10.3389/fnins.2022.752871

Travers, B. G., Powell, P. S., Klinger, L. G., & Klinger, M. R. (2013). Motor Difficulties in Autism Spectrum Disorder: Linking Symptom Severity and Postural Stability. *Journal of Autism and Developmental Disorders*, *43*(7), 1568–1583. https://doi.org/10.1007/s10803-012-1702-x

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

Walsh, J. A., Creighton, S. E., & Rutherford, M. D. (2016). Emotion Perception or Social Cognitive Complexity: What Drives Face Processing Deficits in Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, *46*(2), 615–623. https://doi.org/10.1007/s10803-015-2606-3

Wang, X., Xing, W., & Laffey, J. M. (2018). Autistic youth in 3D game-based collaborative virtual learning: Associating avatar interaction patterns with embodied social presence. *BRITISH JOURNAL OF EDUCATIONAL TECHNOLOGY*, *49*(4, SI), 742–760. https://doi.org/10.1111/bjet.12646

Wei, Q., Cao, H., Shi, Y., Xu, X., & Li, T. (2023). Machine learning based on eye-tracking data to identify Autism Spectrum Disorder: A systematic review and meta-analysis. *Journal of Biomedical Informatics*, *137*, 104254. https://doi.org/10.1016/j.jbi.2022.104254

Wen, T. H., Cheng, A., Andreason, C., Zahiri, J., Xiao, Y., Xu, R., Bao, B., Courchesne, E., Barnes, C. C., Arias, S. J., & Pierce, K. (2022). Large scale validation of an early-age eye-tracking biomarker of an autism spectrum disorder subtype. *Scientific Reports*, *12*(1), 4253. https://doi.org/10.1038/s41598-022-08102-6

Wieckowski, A. T., & White, S. W. (2017). Application of technology to social communication impairment in childhood and adolescence. *Neuroscience and Biobehavioral Reviews*, *74*, 98–114. https://doi.org/10.1016/j.neubiorev.2016.12.030

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine*, *11*(1), 115–129. https://doi.org/10.1017/S0033291700053332

World Health Organization (WHO). (2019). *International Classification of Diseases, Eleventh Revision (ICD-11)* (11th revis). https://icd.who.int/browse11

Wozniak, R. H., Leezenbaum, N. B., Northrup, J. B., West, K. L., & Iverson, J. M. (2017). The development of autism spectrum disorders: variability and causal complexity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1–2), e1426. https://doi.org/10.1002/wcs.1426

Yeung, M. K. (2022). A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: The specificity of deficits and the role of task characteristics. *Neuroscience & Biobehavioral Reviews*, *133*, 104518. https://doi.org/10.1016/j.neubiorev.2021.104518

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

Zhao, H., Zheng, Z., Swanson, A., Weitlauf, A., Warren, Z., & Sarkar, N. (2018). Design of a haptic-gripper virtual reality system (HG) for analyzing fine motor behaviors in children with autism. *ACM Transactions on Accessible Computing*, *11*(4), 1–21. https://doi.org/10.1145/3231938

Zhao, Z., Tang, H., Alviar, C., Kello, C. T., Zhang, X., Hu, X., Qu, X., & Lu, J. (2022). Excessive and less complex body movement in children with autism during face-to-face conversation: An objective approach to behavioral quantification. *Autism Research*, *15*(2), 305–316. https://doi.org/10.1002/aur.2646

Zhao, Z., Zhu, Z., Zhang, X., Tang, H., Xing, J., Hu, X., Lu, J., Peng, Q., & Qu, X. (2021). Atypical Head Movement during Face-to-Face Interaction in Children with Autism Spectrum Disorder. *Autism Research*, *14*(6), 1197–1208. https://doi.org/10.1002/aur.2478

Zhao, Z., Zhu, Z., Zhang, X., Tang, H., Xing, J., Hu, X., Lu, J., & Qu, X. (2022). Identifying Autism with Head Movement Features by Implementing Machine Learning Algorithms. *Journal of Autism and Developmental Disorders*, *52*(7), 3038–3049. https://doi.org/10.1007/s10803-021-05179-2

2. Publication I: Reduced stereotypicality and spared use of facial expression predictions for social evaluation in autism

International Journal of Clinical and Health Psychology 24 (2024) 100440

Contents lists available at ScienceDirect

International Journal of Clinical and Health Psychology

journal homepage: www.elsevier.es/ijchp

Reduced stereotypicality and spared use of facial expression predictions for social evaluation in autism

Marta Robles[a,b], Irene Ramos-Grille[a,c], Amaia Hervás[d,e], Enric Duran-Tauleria[e], Jordi Galiano-Landeira[a], Jolie B. Wormwood[f], Christine M. Falter-Wagner[b], Lorena Chanes[a,g,h,*]

[a] Department of Clinical and Health Psychology, Universitat Autònoma de Barcelona, Barcelona, Spain
[b] Department of Psychiatry and Psychotherapy, LMU University Hospital, LMU Munich, Germany
[c] Division of Mental Health, Consorci Sanitari de Terrassa, Terrassa, Catalunya, Spain
[d] Child and Adolescent Mental Health Service, Hospital Universitari Mútua de Terrassa, Barcelona, Spain
[e] Institut Global d'Atenció Integral del Neurodesenvolupament (IGAIN), Barcelona, Spain
[f] University of New Hampshire, Durham, NH, USA
[g] Institut de Neurociències, Universitat Autònoma de Barcelona, Barcelona, Spain
[h] Serra Húnter Programme, Generalitat de Catalunya, Barcelona, Spain

ARTICLE INFO

Keywords:
Autism
Emotion
Predictive processing
Facial expressions
Social perception

ABSTRACT

Background/Objective: Autism has been investigated through traditional emotion recognition paradigms, merely investigating accuracy, thereby constraining how potential differences across autistic and control individuals may be observed, identified, and described. Moreover, the *use* of emotional facial expression information for social functioning in autism is of relevance to provide a deeper understanding of the condition.
Method: Adult autistic individuals ($n = 34$) and adult control individuals ($n = 34$) were assessed with a social perception behavioral paradigm exploring facial expression predictions and their impact on social evaluation.
Results: Autistic individuals held less stereotypical predictions than controls. Importantly, despite such differences in predictions, the *use* of such predictions for social evaluation did not differ significantly between groups, as autistic individuals relied on their predictions to evaluate others to the same extent as controls.
Conclusions: These results help to understand how autistic individuals perceive social stimuli and evaluate others, revealing a deviation from stereotypicality beyond which social evaluation strategies may be intact.

Introduction

Autism has been traditionally described as a neurodevelopmental disorder characterized by symptoms in communication and social interaction, as well as the presence of restricted and repetitive patterns of behavior, interests, or activities (Autism Spectrum Disorder; American Psychiatric Association, APA, 2022). Among other aspects related to emotion and social cognition, facial expression perception in autism has been extensively assessed. Studies have reported primarily deficits, difficulties, or poorer performance on facial expression recognition among autistic individuals (for meta-analyses see Lozier et al., 2014; Uljarevic & Hamilton, 2013; for a recent review and meta-analysis, see Yeung, 2022). Such observations, described in terms of deficits or difficulties, are tightly linked to assumptions about "accuracy" in emotion perception; studies typically assess an individuals' ability to identify stereotypical or prototypical expressions of emotion (e.g., matching a posed, often exaggerated facial expression to one of a limited set of provided emotion words). Importantly, although autistic individuals tend to perform more poorly on such tasks than non-autistic individuals, it is still unclear *why* these differences emerge or what such measures can actually reveal about emotion perception ability in the real world at all.

Much research on emotion perception, including but not limited to research on autism, relies on assumptions about emotion stemming from what is sometimes referred to as the classical view of emotion (for an overview, see Tracy & Randles, 2011). From this perspective, a core

* Corresponding author at: Department of Clinical and Health Psychology, Universitat Autònoma de Barcelona, Carrer de la Fortuna – Edifici B, Campus de la UAB – 08193 Bellaterra, Cerdanyola del Vallès, Barcelona, Spain.
E-mail address: lorena.chanes@uab.cat (L. Chanes).

https://doi.org/10.1016/j.ijchp.2024.100440
Received 15 September 2023; Accepted 8 January 2024
Available online 23 February 2024

facial configuration ("expression") exists for each of a series of emotions that are considered biologically wired and universally shared (for a review, see, e.g., Barrett et al., 2019). Under this view, sadness is always expressed by exhibiting a pouting facial expression, happiness by a smiling facial expression, and so on. Thus, a person's ability to match a posed, stereotypical expression (i.e., a facial expression that is normatively associated with a given emotion category in a particular culture) to a given emotion label may be considered a measure of emotion perception accuracy from this perspective, because it is presumed that there is an objectively correct answer and that the process is consistent with how emotion perception unfolds in everyday life.

However, there is now substantial empirical data challenging these assumptions, as evidence shows that emotions are experienced and expressed in highly variable ways, and that they vary significantly across cultures and situations (Barrett et al., 2019). In particular, facial expressions of emotion are highly variable (see Fernández-Dols & Crivelli, 2013) and individuals fail to consistently make the stereotypical expressions used in emotion perception research at times when they are actually experiencing those emotions (Durán & Fernández-Dols, 2021). These findings are consistent with constructionist views of emotion, such as the Theory of Constructed Emotion (Barret, 2017), which posit that emotions are not biologically discrete but instead represent categories constructed by applying conceptual knowledge about emotion (learned through one's culture, language, and experience) to make meaning of ongoing sensory information arising from the body (e.g., changes in heart rate or breathing). From this perspective, emotions represent heterogeneous categories of instances, such that there is both variability within emotion categories (i.e., not all instances of anger are identical, even for the same person) and similarity across categories (i.e., instances of anger and happiness might both involve smiling or an elevated heart rate). Thus, according to constructionist theories of emotion, variability in the components of emotion, including facial expressions, are expected, undermining the assumption made in most emotion perception research that there is a one-to-one correspondence between a specific facial configuration and the experience of a given emotion (Barrett et al., 2019). Simply put, people do not always pout when sad or smile when happy. As such, typical emotion perception tasks that ask individuals to match posed facial expressions to a given emotion word cannot assess 'accuracy' in emotion perception. Instead, they measure the person's ability to identify normative or stereotypical expressions that belong to a given emotion category in a given culture. Here, we move away from assessing 'accuracy' in emotion perception to examine the extent to which autistic individuals expect to see stereotypical facial expressions in different emotional evocative scenarios.

Indeed, amongst constructivist views, expectations or predictions are posited to be the basis of emotion perception according to The Theory of Constructed Emotion (Barret, 2017), which builds on predictive processing accounts of perception and action, suggesting that emotion experience and perception unfold predictively, with prior experience and beliefs guiding experience and perception in critical ways (Barrett, 2017; Hoemann et al., 2020). In recent years, multiple brain-related conditions, including autism, have been described in terms of predictive processing or predictive coding (for a recent review, see, Smith et al., 2021). The core idea of predictive processing is that the brain constantly issues predictions about what will happen next based on previous experiences, which are then compared to actual sensory input from the body and the external world (see, e.g., Friston, 2005, 2010). When the difference between predictions and actual sensory input, called prediction error, is minimal, *predictions* are thought to drive perception and behavior. Only when the prediction error passes some threshold does actual sensory input more strongly inform perception and behavior, and the model used to issue predictions is updated to reduce such prediction error in the future. Moreover, reliance on predictions and prediction errors are weighted, i.e., are prioritized or not, based on their perceived reliability (i.e., precision; Yon & Frith, 2021). With regards to autism, several models have been proposed, roughly suggesting a privileged processing of prediction error compared to predictions (Friston et al., 2013; Lawson et al., 2014; Pellicano & Burr, 2012; Quattrocki & Friston, 2014a; Sinha et al., 2014; van de Cruys et al., 2014; for a recent review, see, Cannon et al., 2021), leading to a model that may become too complex too soon, including details that may remain overlooked by non-autistic individuals. Practically speaking, prioritizing prediction error could lead autistic individuals to have difficulties generalizing and adapting to the constantly changing world we live in, requiring instead anticipation, sameness, and routines, a core feature of autism (American Psychiatric Association, 2022). In the context of emotion perception, an overreliance on prediction error would be consistent with expectations that are more nuanced and less normative.

In the present study, we use a social perception task in which facial expression predictions are induced through short written scenarios and subsequently confirmed or violated to assess how much prediction error a person experiences and how normative or stereotypical their predictions about facial expressions of emotion are. In the task, participants are presented with a picture of a person and read a text that describes a situation (scenario) thought to be normatively associated with a specific emotion (e.g., attending a funeral and sadness; enjoying time with friends and happiness). We ask them to imagine how the person would look in that scenario, and then show them the person's face depicting a stereotypical emotion expression that either matches or mismatches the emotion normatively associated with the scenario. Subsequently, we assess how much the face matched their prediction about what the person would look like and how much they liked the person.

In previous studies using this task, healthy adult participants were shown to hold fairly stereotypical predictions about facial expressions in emotion contexts (Chanes et al., 2018): participants rated stereotypical facial expressions matching the emotion evoked by the scenario (e.g., a smiling face for a happy scenario) as more similar to what they had imagined than when the stereotypical facial expression did not match the emotion evoked by the preceding scenario (e.g., a smiling face for a sad scenario). Critically, faces posing stereotypical expressions that were rated as more predictable (i.e., as more in line with the perceiver's expectations) were also judged as more likable and trustworthy. Thus, in emotion contexts, facial expression stereotypes seem to drive individuals' predictions about what expression they will encounter in a given context, and people are judged more favorably when their expression matches a perceiver's predictions about how they should look (i.e., the perceiver's internal model).

Using this same task, we here explore facial expression perception in adult autistic individuals, assessing the extent to which they predict to encounter stereotypical facial expressions in emotionally evocative contexts. By assessing the impact of such expectations on likability ratings, we aimed to go one step further and study whether these predictions impact social evaluation in autism, which may contribute to better understand social function in this condition. We hypothesized that predictions about facial expressions would be less stereotypical for autistic than control individuals, and that such differences could be associated with differences in social evaluation.

Materials and methods

Participants

Thirty-four autistic adult individuals (sixteen females, eighteen males; autistic group) and 34 adult individuals without a current or history of a diagnosis of any psychiatric or neurological condition and not taking psychoactive medication (sixteen females, eighteen males; control group) participated in the study. Sample size was based on previous studies with overall-healthy individuals (Chanes et al., 2018) and individuals with depression (Ramos-Grille et al., 2022), and in line with studies investigating facial expression perception in autism (see, e.g Yeung, 2022). Autistic individuals were recruited from the Global

Institute of Neurodevelopment Integrated Care (IGAIN), a healthcare center specialized in autism located in Barcelona (Spain). The diagnosis of autism spectrum disorder (ASD), as well as the absence of an intellectual developmental disorder, was confirmed by clinical experts at the center according to DSM-5 (APA, 2013) through extensive clinical evaluation. Inclusion criteria for the autistic group were: (i) a confirmed diagnosis of ASD, (ii) no history of brain injury or other neurological conditions, (iii) no presence of an intellectual developmental disorder, (iv) age $\geq$ 18 years old, (v) normal or corrected-to-normal vision, and (vi) native Spanish speaker or bilingual Catalan-Spanish. In this group, twenty-one participants had one or more comorbid diagnoses. More specifically, 13 of them presented one comorbid diagnosis and 8 presented two or more comorbid diagnoses, roughly reflecting the ratios of comorbid diagnoses in autism observed in previous literature (APA, 2022) (see Table 1 for more details). Twenty-one participants were taking one or more medications.

Control individuals were recruited in Barcelona area through word-of-mouth and advertisements shared on social media. They were selected to match autistic participants' gender and age within a range of $\pm$ 5 years. For control individuals, the same inclusion criteria applied, as well as not having been diagnosed with any psychiatric or neurological condition and not be taking psychoactive medication. One control participant was excluded after data collection but before any data analysis, because of a diagnosed psychiatric condition reported at the end of the session. This participant was replaced by another participant of the same gender and similar age before any data analysis.

The study was approved by the Autonomous University of Barcelona's Institutional Review Board. All participants gave written informed consent and did not receive any monetary compensation for their participation. There was no community involvement in the reported study.

The autistic and control groups were 28 $\pm$ 2 years old (*Mean* $\pm$ *SE*) and 31 $\pm$ 1 years old (*Mean* $\pm$ *SE*), respectively, the difference being small although significant (Mann-Whitney two-tailed U-test, $U = 754.5$, $p = 0.031$, $rB = 0.305$, 95 % CI [0.039, 0.531]) (Table 1).

We performed an additional analysis on control participants focusing on autistic traits. For that specific analysis we included not only control participants but also 7 control participants from a different ongoing study (with the same inclusion criteria and study characteristics), adding up to a total of 41 individuals (Autism-like traits group; see Table 1 for more details). This group was 30 $\pm$ 1 years old (*Mean* $\pm$ *SE*).

Experimental procedure

Data collection took place at the healthcare center (IGAIN). First, sociodemographic data was collected (reported age, sex, gender, handedness, and education level). Next, participants performed a predictive processing and social perception task (Chanes et al., 2018) after which they completed the abridged Version of the Autism-Spectrum Quotient, (AQ-Short; Hoekstra et al., 2011; Spanish version used: Lugo-Marín et al., 2019) among other questionnaires not used for this study (Positive and Negative Affect Scale, Watson et al., 1988; Emotion Regulation Questionnaire, Gross & John, 2003; short version of the UPPS-P impulsive behavior scale, Billieux et al., 2012). Diagnoses (including co-occurring conditions), and medications at the time of study participation for the autistic group were collected from medical records.

Predictive processing and social perception task

Stimuli were presented using E-Prime 3 (Psychology Software Tools, Inc., Pittsburgh, PA, USA) running on an HP ProBook 640 G4 (display size: 14'', resolution: 1920 × 1080).

Table 1
Sociodemographic and clinical variables for the autistic ($n = 34$) and control ($n = 34$) groups, as well as for the autism-like traits group ($n = 41$).

Variables	Autistic group		Control group		Autism-like traits group	
	n/mean	%/[SD]	n/mean	%/[SD]	n/mean	%/[SD]
GENDER						
Women	16	53 %	16	53 %	19	46 %
Men	18	47 %	18	47 %	22	54 %
AGE (years)	28	[10]	31	[7]	30	[7]
EDUCATION LEVEL						
Primary or lower	1	3 %	0	0 %	0	0 %
High school	28	82 %	13	38 %	14	34 %
University	5	15 %	21	62 %	27	66 %
AQ-SHORT SCORE	72	[11]	53	[10]	54	[11]
COMORBID DIAGNOSES	21	62 %				
One comorbid diagnosis	13	38 %				
Two or more comorbid diagnoses	8	24 %				
Anxiety Disorders	17	50 %				
Generalized Anxiety Disorder	8	24 %				
Social Anxiety	9	26 %				
Depressive Disorders	4	12 %				
Major Depression	2	6 %				
Unspecified Depressive Disorder	2	6 %				
Neurodevelopmental Disorders	6	18 %				
Attention Deficit Hyperactivity Disorder	5	15 %				
Specific Learning Disorder (Dyscalculia)	1	3 %				
Other disorders						
Unspecified Bipolar Disorder	1	3 %				
Obsessive-Compulsive Disorder	2	6 %				
Post-Traumatic Stress Disorder	1	3 %				
PSYCHOLOGICAL TREATMENT	34	100 %				
PHARMACOLOGICAL TREATMENT	21	62 %				
Alpha-2 adrenergic agonist	7	21 %				
Benzodiazepines	3	9 %				
Central nervous system stimulant	1	3 %				
Norepinephrine-dopamine reuptake inhibitor	13	38 %				
Selective serotonin reuptake inhibitor	7	21 %				
Serotonin-norepinephrine reuptake inhibitors	6	18 %				
Tetracyclic antidepressants	2	6 %				

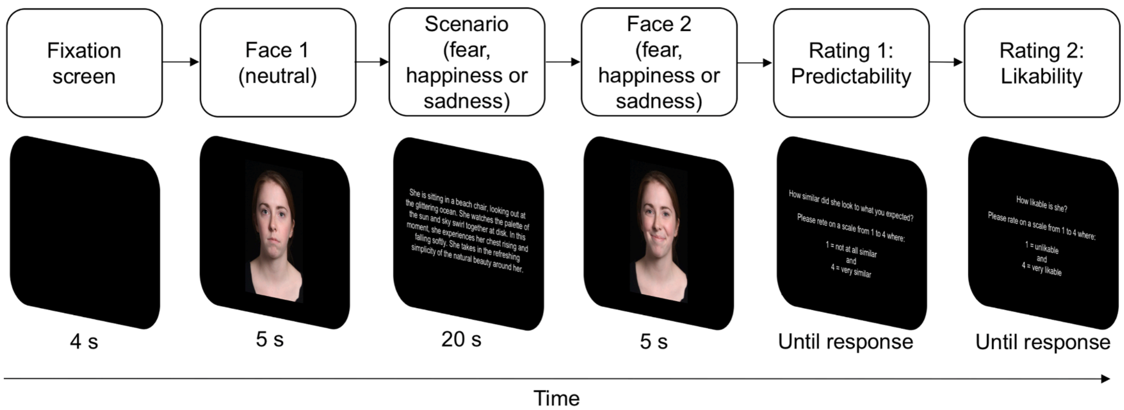

Fig. 1. Sequence of screens on a sample experimental trial. After a fixation screen, a target person displaying a neutral facial expression was presented. Next, a scenario (short text) was displayed evoking either fear, happiness, or sadness. Then, the same target person was shown exhibiting a stereotypical facial expression for fear, happiness, or sadness. The facial expression could match the stereotypical expression for the emotion evoked by the scenario (matched trials) or not (nonmatched trials). Afterwards, two ratings (predictability and likability) were performed.

Each trial began with a black fixation screen (4 s) (Fig. 1). A photograph of a target person with a neutral facial expression was then shown at the center of the screen (Face 1; 5 s). Next, a short story (Scenario; 20 s) was displayed in white font. Each scenario, describing a situation experienced by the target person, was aimed to evoke one of the three following emotions: fear, happiness, or sadness. Participants were asked to imagine, while reading through the story, how the target person would look in that scenario. After the scenario, a second photograph of the target person was displayed, exhibiting a stereotypical facial expression for one of the three possible evoked emotions (fear, happiness, or sadness; Face 2; 5 s). On some trials, the stereotypical expression corresponded to the emotion evoked by the scenario (matched trials; e.g., a pouting facial expression following a scenario evoking sadness) and on some trials it did not (nonmatched trials; e.g., a pouting facial expression following a scenario evoking happiness or fear). Finally, participants were asked to complete two ratings on four-point scales with no time limit. In the first rating, they were asked to indicate how similar the target person looked (Face 2) compared to what they had imagined while reading the scenario, on a scale from 1 = *"not at all similar"* to 4 = *"very similar"* (predictability rating; Rating 1). In the second rating, participants indicated how likable the target person was, on a scale from 1 = *"very unlikable"* to 4 = *"very likable"* (likability rating; Rating 2). Thus, the task was designed to tackle subjective predictions about facial configurations. Emotion categories were not explicitly mentioned to participants at any time. We aimed to assess how well the images presented matched participants' predictions, which may be considered as a rough behavioral correlate of prediction error. Because there were no right/wrong answers, no accuracy was computed, and results may not be described in terms of a better/worse performance.

Prior to the start of the task, participants were asked to read task instructions displayed on the screen, which included sample photographs and sample rating screens. After going through the instructions and asking the researcher any questions they might had, participants were asked to verbally summarize the task to the researcher. Once the researcher confirmed that the task had been correctly understood by participants, they performed 3 practice trials. These trials were all matched trials (one evoking each of the three possible emotions: fear, happiness, and sadness). The researcher observed participants' responses to these trials. If the predictability rating on any of these three trials was relatively low (1 or 2), the researcher asked questions to the participant to confirm that the low ratings on those trials were not due to a lack of understanding of the instructions, a lack of understanding of the texts, or any other aspect suggesting a lack of capability to perform the task. The script followed by the researcher was: "I have seen that you have responded 1 or 2 on some/all of the trials. How did you imagine/expect that the person would look?". By the participant's response, the researcher would confirm that the low rating was due to differences in their predictions with regards to the displayed photograph, rather than a lack of understanding of the instructions, the texts or other aspects preventing the participant from performing the task adequately.

The task included 45 experimental trials divided into 5 blocks. Each block consisted of 9 trials and participants were offered to take a short break in between if they wished. The 45 experimental trials included 15 scenarios evoking each of the three emotion categories: fear, happiness, and sadness. Among the 15 scenarios evoking a given emotion, 9 scenarios were followed by the corresponding stereotypical facial expression (matched trials). Six scenarios were followed by a stereotypical facial expression for a different emotion than the one evoked by the scenario, 3 for each of the other two emotion categories (nonmatched trials). This added up to a total of 27 matched trials and 18 nonmatched trials per participant. Identities of the models shown for each scenario, scenario order, and match condition assigned to a given scenario were all pseudorandomized within participants.

A different target person (identity) was used for each of the 48 trials (3 practice trials: 2 females, 1 male; 45 experimental trials: 28 females, 17 males). Color photographs (400 × 600 pixels) of human faces with closed mouths and a direct gaze, which belong to the Interdisciplinary Affective Science Laboratory[1] (www.affective-science.org), were used (for more details, see Chanes et al., 2018). We used the scenarios used in previous studies (Chanes et al., 2018, originally developed by Wilson--Mendenhall et al., 2013), translated into Spanish (used in Draganov et al., 2023; Ramos-Grille et al., 2022). The scenarios narrated a scene matching the target person's gender, and evoked either happiness, sadness, or fear, with high or low arousal for each emotion (e.g., "She is sitting in a beach chair, looking out at the glittering ocean. She watches the palette of the sun and sky swirl together at dusk. In this moment, she experiences her chest rising and falling softly. She takes in the refreshing simplicity of the natural beauty around her".).

Questionnaire

We used the Spanish abridged version of the Autism-Spectrum Quotient, (AQ-Short; Hoekstra et al., 2011; Spanish version used: Lugo-Marín et al., 2019) in order to measure the presence of autistic

[1] Development of the Interdisciplinary Affective Science Laboratory (IASLab) Face Set was supported by the National Institutes of Health Director's Pioneer Award (DP1OD003312) to Lisa Feldman Barrett. More information is available online at www.affective-science.org.

traits across participants. The AQ-Short consists of 28 self-reported items describing typical autistic traits such as "I find it difficult to make new friends", with responses on a 4-point scale, with possible answers being "1 = definitely agree'', "2 = slightly agree'', "3 = slightly disagree'', and "4 = definitely disagree''. Scoring is reversed for items in which an "agree'' response is not characteristic of autism and item scores are summed. Lugo-Marín et al. (2019) propose a cut-off score of 65 to identify potentially autistic individuals with a sensitivity of 0.98 and a specificity of 0.84. The results of the reliability analysis suggest a good internal structure (intraclass correlation coefficients ranging from 0.90 to 0.97), in line with the results found for the original version (Hoekstra et al., 2011).

Data analysis

Non-aggregated data from individual trials were analyzed using hierarchical linear modeling (HLM 7.0; Scientific Software International, Inc., Skokie, IL, USA). HLM analysis was used to avoid aggregation across trials and model variability in trial-by-trial performance nested within each participant. We used a continuous sampling model with participants treated as a random factor, and a restricted maximum likelihood method of estimation for model parameters (Raudenbush & Bryk, 2002). Continuous trial-level variables (e.g., predictability ratings) were centered around each participant's mean when entered as predictors in the models. Dummy-coded variables (e.g., match condition) were uncentered (Enders & Tofighi, 2007). All HLM models had random intercepts. Additionally, when analyzing traits rather than autism as a category, a linear regression model was computed. The linear regression and additional statistical analyses were carried out using JASP (JASP Team, 2022; version 0.9.2). For comparisons across control and autistic groups, we used t-tests or Mann-Whitney *U*-tests when normality assumptions were not met. Prism 9 (GraphPad Software, San Diego, CA, USA) was used for data visualization.

Results

Self-reported autistic traits for autistic and control groups

As expected, AQ-Short scores differed significantly across groups (Mann-Whitney two-tailed *U*-test, $U = 110$, $p < 0.001$, $rB = -0.840$, 95 % *CI* [0.883, -0.681]), the autistic group presenting more autistic traits (*Mean* ± *SE*: 72 ± 2) than the control group (*Mean* ± *SE*: 53 ± 2) (Table 1). AQ-Short scores from one autistic individual were missing due to tiredness and inability to finalize responses to the questionnaire. The autism-like traits group had AQ-Short scores of 54 ± 2 (*Mean* ± *SE*).

Predictability of facial expressions of emotion

Overall

An HLM analysis was computed, with trial-level predictability ratings as the outcome variable and group (autistic and control) as a dummy-coded subject-level predictor variable. This analysis revealed that, overall, autistic individuals rated facial expressions as significantly less similar to what they expected (lower predictability ratings) (*Mean* ± *SE*: 2.49 ± 0.07) than control individuals (Mean ± *SE*: 2.65 ± 0.04) ($B = -0.17$, $SE = 0.08$, $t(66) = 2.16$, $p = 0.034$). This is consistent with autistic individuals having increased prediction error relative to controls.

By match condition

Next, we conducted an HLM analysis with trial-level predictability ratings as the outcome variable, match condition as a dummy-coded trial-level predictor variable, and group (autistic and control) as a dummy-coded subject-level predictor variable, with the two-way interaction term included. As expected, matched facial expressions (stereotypical facial expressions for the evoked emotion) were rated as significantly more predictable (*Mean* ± *SE*: 3.25 ± 0.05) than nonmatched facial expressions (*Mean* ± *SE*: 1.76 ± 0.05) for the control group ($B = 1.49$, $SE = 0.06$, $t(66) = 24.85$, $p < 0.001$), which is consistent with prior studies (Chanes et al., 2018; Draganov et al., 2023; Ramos-Grille et al., 2022). Autistic individuals also rated matched facial expressions as significantly more predictable (*Mean* ± *SE*: 2.93 ± 0.09) than nonmatched (*Mean* ± *SE*: 1.82 ± 0.08) facial expressions ($B = 1.11$, $SE = 0.10$, $t(66) = 11.05$, $p < 0.001$). However, a significant interaction indicated that the difference in predictability ratings for matched vs. nonmatched facial expressions (namely 'match effect'), was significantly less pronounced among autistic individuals than among control individuals ($B = -0.39$, $SE = 0.12$, $t(66) = 3.30$, $p = 0.002$).

To explore this interaction further, we separately analyzed predictability ratings for trials with matched and nonmatched facial expressions. These analyses revealed that autistic individuals rated the expressions as significantly less predictable than control individuals on trials where the facial expression matched the stereotypical expression for the evoked emotion (matched trials; $B = 0.32$, $SE = 0.10$, $t(66) = 3.25$, $p = 0.002$), but there were no differences in predictability ratings on trials where the facial expression did not match the stereotypical expression for the evoked emotion (nonmatched trials; $B = -0.06$, $SE = 0.09$, $t(66) = 0.69$, $p = 0.490$). This pattern of results suggests that autistic individuals have less stereotypical predictions for facial expressions of emotion than those in the control group; they found stereotypical expressions matched to the emotion evoked by the scenario (i.e., smiling faces following a normatively happy scenario) less in line with their own predictions for what the target person would look like in that scenario than did controls.

By match and emotion condition

We further explored whether the observed differences between autistic individuals vs. control individuals in terms of predictability ratings were consistent across the three different emotion categories evoked by the scenarios (see Fig. 2). To do so, we conducted the above analyses separately for trials with each of the three different evoked emotions (happiness, sadness, and fear). These analyses revealed that, compared to controls, autistic individuals exhibited a less pronounced difference in predictability ratings on matched vs. nonmatched trials for scenarios evoking happiness (autistic individuals (*Mean* ± *SE*): 3.28 ± 0.09 vs. 1.53 ± 0.08; control individuals (*Mean* ± *SE*): 3.54 ± 0.05 vs. 1.36 ± 0.07; $B = -0.43$, $SE = 0.15$, $t(66) = 2.87$, $p = 0.006$), sadness (autistic individuals (*Mean* ± *SE*): 2.75 ± 0.10 vs. 1.97 ± 0.09; control individuals (*Mean* ± *SE*): 3.18 ± 0.07 vs. 1.93 ± 0.07; $B = -0.48$, $SE = 0.13$, $t(66) = 3.71$, $p < 0.001$), and fear (autistic individuals (*Mean* ± *SE*): 2.76 ± 0.10 vs. 1.97 ± 0.10; control individuals (*Mean* ± *SE*): 3.03 ± 0.07 vs. 1.99 ± 0.07; $B = -0.25$, $SE = 0.15$, $t(66) = 1.61$, $p = 0.112$), though this interaction only reached significance for scenarios evoking happiness and sadness. Further, analyses revealed that autistic individuals rated matched expressions as significantly less predictable than controls for all three emotion scenarios (happiness: $B = 0.26$, $SE = 0.10$, $t(66) = 2.46$, $p = 0.016$; sadness: $B = 0.44$, $SE = 0.13$, $t(66) = 3.49$, $p < 0.001$; fear: $B = 0.27$, $SE = 0.12$, $t(66) = 2.25$, $p = 0.028$). No differences in predictability ratings between autistic individuals and control participants were observed for nonmatched trials for any emotion condition (happiness: $B = -0.17$, $SE = 0.10$, $t(66) = 1.64$, $p = 0.105$; sadness: $B = -0.04$, $SE = 0.11$, $t(66) = 0.36$, $p = 0.722$; fear: $B = 0.02$, $SE = 0.12$, $t(66) = 0.17$, $p = 0.869$). Taken together, these results indicate that the effects observed do not seem to depend on the specific evoked emotion but rather emerge across emotion categories.

Facial expression predictions with regards to autism-like traits

We explored if the above-described reduced match effect (i.e., difference in predictability ratings for matched vs. nonmatched facial expressions) for autistic vs. control individuals is also observed when considering autism-like traits in non-autistic individuals. Only for this additional analysis, data from our control sample ($N = 34$) as well as

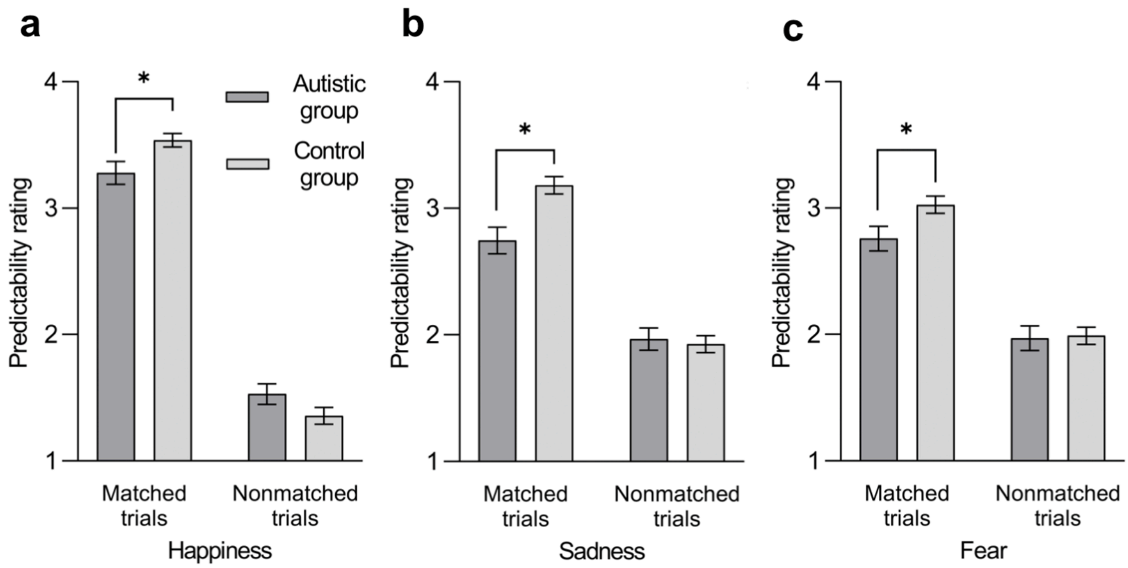

Fig. 2. Predictability ratings for autistic and control individuals by scenario evoked emotion and match condition. Means and standard errors for the autistic (dark gray) and the control (light gray) groups. Match effect (difference between ratings on matched and nonmatched trials) was reduced in autistic vs. control individuals. This effect was driven primarily by differences in predictability ratings on matched trials. Asterisks indicate $p < 0.05$.

from 7 further control participants from a different ongoing study were used (total $N = 41$). A match effect variable was computed for each participant by subtracting the mean of each participant's predictability ratings on matched trials and nonmatched trials. We then performed a linear regression with AQ-Short scores as the predictor variable and match effect as the outcome variable. This analysis revealed that, for non-autistic individuals, higher autism-like traits were associated with a smaller match effect ($R^2 = 0.109$; $F(1, 39) = 4.767$, $t = -2.183$ p $=$ 0.035). Thus, the association between autism and having less stereotypical predictions is observed both when comparing those with and without an autism diagnosis, as well as for non-autistic individuals with more (vs. fewer) autism-like traits.

Predictability of facial expressions and social evaluation

The impact of facial expression predictability on social evaluation was assessed as the relationship between predictability and likability ratings using an HLM analysis with trial-level likability ratings as the outcome variable, predictability ratings as a trial-level predictor variable, and group (autistic and control individuals) as a dummy-coded subject-level predictor variable, with the two-way interaction term included in the model. Consistent with prior work (Chanes et al., 2018), this analysis revealed a positive relationship between predictability and likability ratings across conditions. Facial expressions that were rated as more predictable were also rated as more likable for both groups (control individuals: $B = 0.23$, $SE = 0.03$, $t(66) = 6.70$, $p < 0.001$; autistic individuals: $B = 0.25$, $SE = 0.04$, $t(66) = 6.50$, $p < 0.001$). Importantly, this relationship did not differ significantly between groups ($B = -0.01$, $SE = 0.05$, $t(66) = 0.25$, $p = 0.801$). This pattern of results suggests that individuals like others more when they display predicted expressions (i.e., expressions that better match the perceiver's own prediction of what someone will look like in a given scenario), and this relationship between predictability and liking does not differ across autistic individuals and control individuals.

Discussion

Compared to controls, autistic individuals reported greater inconsistencies between stereotypical facial expressions of emotion and their own predictions about what someone's facial expression would be in a given context, suggesting their predictions are less stereotypical than controls'. Similarly to controls, however, autistic individuals liked others more when their facial expression better matched their own predictions about what it would look like in a given scenario.

Although the autistic group had lower predictability ratings than the control group overall, i.e., a larger perceived difference between what they expected and the actual stimuli displayed, this effect was driven by ratings in the matched condition only; rather than being systematically lower across all trials, predictability ratings for autistic individuals were lower specifically for matched trials, where the stereotypical facial expression for the emotion depicted by the scenario was displayed. There were no differences observed between the autistic and control groups in terms of predictability ratings, for nonmatched trials, where a stereotypical facial expression for a different emotion than that depicted by the scenario was displayed. That we observed similar predictability ratings for nonmatched trials across groups indicates that both groups found categorically normative mismatches (i.e., a smiling face in a sad scenario) unexpected to the same extent. Thus, findings point to autistic individuals having less stereotypical predictions when thinking about what someone's facial expression will be in a given emotionally evocative situation. An alternative possibility is that autistic individuals have a general "deficit", i.e., did not have clear facial expression predictions in emotion contexts, though this explanation seems less likely given our rigorous protocol for ensuring understanding and compliance with task instructions and ability to perform it among all participants. Critically, autistic individuals also had significantly higher predictability ratings in matched than nonmatched trials (although this difference was smaller among autistic participants than controls), suggesting they were making predictions to some extent as otherwise predictability ratings would be expected not to differ significantly across both types of trials within the autistic group. Moreover, we found converging evidence looking at autism-like traits (as assessed by the AQ-Short) only in non-autistic individuals: individuals that showed higher characteristics of autism also showed a smaller difference in predictability ratings between matched and nonmatched trials. Furthermore, autistic individuals relied on their predictions to the same extent than controls to rate likability. Nevertheless, we did not directly assess the content of participants' predictions nor their confidence in them, so future work should examine the extent to which these qualities of the predictions differ across autistic and control participants in general as well as in stereotypical contexts specifically.

This lack of stereotypical predictions for facial expressions of emotion among autistic individuals is consistent with observations of lower “accuracy” for autistic vs. control individuals in previous studies using traditional emotion recognition tasks framed in classical views of emotion perception (for meta-analyses see Lozier et al., 2014; Uljarevic & Hamilton, 2013; Yeung, 2022). In these tasks, the facial expression stipulated as correct for a given emotion category is typically a highly stereotypical expression for that emotion. To the extent that autistic individuals do not expect expressions as stereotypical as non-autistic individuals, this could lead to less accurate matching of expressions and emotion words in a typical emotion perception task.

Our findings are also consistent with recent predictive coding accounts of autism that suggest a privileged processing of prediction error in autistic individuals (for a recent review, see Cannon et al., 2021). This overreliance on prediction error may lead autistic individuals to incorporate many detailed circumstances and features into their predictions (i.e., internal model), leading to their expectations of facial expressions of emotion being more detailed and less generalizable or stereotypical than non-autistic individuals'.

In the present study we aimed to go beyond assessing facial expression predictions, additionally assessing the relevance of such predictions for social evaluation (likability ratings). In general, individuals tend to like people more when they display facial expressions that better match their predictions for a given emotionally evocative context (Chanes et al., 2018). Importantly, no differences were observed across groups in terms of the strength of this association, with autistic individuals using facial expression predictions to assess likability to the same extent as control individuals. These results show similarities with those of a recent study that found differences between autistic and control individuals when looking at accuracy levels in a facial emotion recognition task, but not when looking at a social functioning-related aspect. More specifically, while autistic individuals exhibited a less accurate retrospective emotion identification, a related social domain (specifically the authors evaluated empathy) appeared to be intact (Santiesteban et al., 2021). Thus, whereas autistic individuals seem to indeed differ from control individuals in their predictions about facial expressions, holding less stereotypical ones, they seem to use their predictions to a similar extent for social evaluation.

Our study is not without limitations. First, the autistic group included participants with co-occurring mental health conditions, and most were taking medication, something that should be taken into account in future studies. At the same time, we deemed it important to test a representative sample of individuals with autism, who do often show co-occurring mental health conditions. Second, all the included participants were adults without an intellectual developmental disorder, although recent studies show that up to a third of the autistic population present differences in intellectual development (Maenner et al., 2021). Thus, future studies may address these processes within a wider range of intellectual capabilities, aiming to better characterize the whole autistic spectrum. Also, in the present study we only asked whether their expectations were similar to the displayed expressions, but we did not assess what their expectations were specifically. Future work may characterize autistic individuals' expectations with regards to controls and assess whether differences exist with regards to clarity, level of detail, or other specific aspects. Moreover, we only assessed three emotion categories (happiness, sadness, and fear) and one social evaluation aspect (likability). Other emotions may be addressed in future studies in order to better explore and characterize the observed effects. Likewise, other ratings may be used in the future in order to explore social evaluation more thoroughly (e.g., trustworthiness, Chanes et al., 2018). Finally, we aimed to assess predictability as a whole, so how different aspects of the facial expression or emotion scenarios (e.g., facial configuration, intensity of the expression, specific cognitive aspects of image perception, narrative interpretations, affective response to narratives, etc.) may have contributed to the present results, remains to be further explored in the future.

Conclusion

The present study provides a new framework to understand how autistic individuals perceive social stimuli and evaluate others, revealing a deviation from stereotypicality in their predictions about others' emotion expressions. Importantly, though, the use of facial expression predictions in social evaluation (i.e., judgments of likability), appears to be intact among autistic individuals.

Funding

LC was funded by the Ministry of Science and Innovation of Spain (PSI2017-88416-R and PID2020-119677RB-I00), the Government of Catalonia (2017 SGR 1612 and SGR 2021 01010) and ICREA Acadèmia. CFW was funded by the Deutsche Forschungsgemeinschaft (FA 876/3-1 and FA 876/5-1).

CRediT authorship contribution statement

Marta Robles: Formal analysis, Investigation, Data curation, Writing – original draft, Project administration. **Irene Ramos-Grille:** Investigation, Writing – review & editing. **Amaia Hervás:** Investigation, Writing – review & editing. **Enric Duran-Tauleria:** Investigation, Resources, Writing – review & editing. **Jordi Galiano-Landeira:** Writing – review & editing, Visualization. **Jolie B. Wormwood:** Writing – review & editing. **Christine M. Falter-Wagner:** Writing – review & editing, Funding acquisition. **Lorena Chanes:** Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgment

We would like to thank the participants of this study and their families, as well as Laia Peran for her help during the participants' recruitment.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (Fifth Edition). American Psychiatric Association. Text Revision (DSM-5-TR).

Barret, L. F. (2017). *How emotions are made: The secret life of the brain*. New York, NY: Houghton Mifflin Harcourt.

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(1), 1–23.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Billieux, J., Rochat, L., Ceschi, G., Carré, A., Offerlin-Meyer, I., Defeldre, A. C., & Van der Linden, M. (2012). Validation of a short French version of the UPPS-P impulsive behavior scale. *Comprehensive Psychiatry, 53*(5), 609–615. https://doi.org/10.1016/j.comppsych.2011.09.001

Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in autism spectrum disorder: A systematic review of empirical evidence. *Autism Research, 14*(4), 604–630. https://doi.org/10.1002/AUR.2482

Chanes, L., Wormwood, J. B., Betz, N., & Barrett, L. F. (2018). Facial expression predictions as drivers of social perception. *Journal of Personality and Social Psychology, 114*(3), 380–396. https://doi.org/10.1037/pspa0000108

Draganov, M., Galiano-Landeira, J., Doruk Camsari, D., Ramírez, J.-E., Robles, M., & Chanes, L. (2023). Noninvasive modulation of predictive coding in humans: Causal evidence for frequency-specific temporal dynamics. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhad127

Durán, J. I., & Fernández-Dols, J.-M. (2021). Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion, 21*(7), 1550–1569. https://doi.org/10.1037/emo0001015 (Washington, D.C.).

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Fernández-Dols, J.-M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review, 5*(1), 24–29. https://doi.org/10.1177/1754073912457229

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*, 127–138. https://doi.org/10.1038/nrn2787

Friston, K. J., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: Comment on pellicano and burr. *Trends in Cognitive Sciences, 17*, 1. https://doi.org/10.1016/j.tics.2012.11.003

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., & van der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of Autism and Developmental Disorders, 41*(5), 589–596. https://doi.org/10.1007/s10803-010-1073-0

Hoemann, K., Devlin, M., & Barrett, L. F. (2020). Comment: Emotions are abstract, conceptual categories that are learned by a predicting brain. *Emotion Review, 12*(4), 253–255.

JASP Team. (2022). JASP (Version 0.9.2) [Computer software].

Lawson, R. P., Rees, G. F. K. J., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience, 8*(MAY). https://doi.org/10.3389/fnhum.2014.00302

Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology, 26*(4 Pt 1), 933–945. https://doi.org/10.1017/S0954579414000479

Lugo-Marín, J., Díez-Villoria, E., Magán-Maganto, M., Pérez-Méndez, L., Alviani, M., de la Fuente-Portero, J. A., & Canal-Bedia, R. (2019). Spanish validation of the autism quotient short form questionnaire for adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*(11), 4375–4389. https://doi.org/10.1007/s10803-019-04127-5

Maenner, M. J., Shaw, K. A., Bakian, A. V, Bilder, D. A., Durkin, M. S., Esler, A., & Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *Morbidity and Mortality Weekly Report, 70*(11), 1–16. https://doi.org/10.15585/mmwr.ss7011a1. Surveillance Summaries (Washington, D.C. : 2002).

Pellicano, E., & Burr, D. (2012). When the world becomes "too real": A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16*, 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Quattrocki, E., & Friston, K. (2014). Autism, oxytocin and interoception. *Neuroscience and Biobehavioral Reviews, 47*, 410–430. https://doi.org/10.1016/j.neubiorev.2014.09.012

Ramos-Grille, I., Weyant, J., Wormwood, J. B., Robles, M., Vallès, V., Camprodon, J. A., & Chanes, L. (2022). Predictive processing in depression: Increased prediction error following negative valence contexts and influence of recent mood-congruent yet irrelevant experiences. *Journal of Affective Disorders, 311*, 8–16. https://doi.org/10.1016/j.jad.2022.05.030

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Santiesteban, I., Gibbard, C., Drucks, H., Clayton, N., Banissy, M. J., & Bird, G. (2021). Individuals with autism share others' emotions: Evidence from the continuous affective rating and empathic responses (CARER) task. *Journal of Autism and Developmental Disorders, 51*(2), 391–404. https://doi.org/10.1007/s10803-020-04535-y

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America, 111*(42), 15220–15225. https://doi.org/10.1073/pnas.1416797111

Smith, R., Badcock, P., & Friston, K. J. (2021). Recent advances in the application of predictive coding and active inference models within clinical neuroscience. *Psychiatry and Clinical Neurosciences, 75*(1), 3–13. https://doi.org/10.1111/pcn.13138

Tracy, J. L., & Randles, D. (2011). Four models of basic emotions: A review of Ekman and Cordaro, izard, Levenson, and Panksepp and watt. *Emotion Review, 3*, 397–405. https://doi.org/10.1177/1754073911410747

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

van de Cruys, S., Evers, K., van der Hallen, R., van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review, 121*(4), 649–675. https://doi.org/10.1037/a0037665

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037//0022-3514.54.6.1063

Wilson-Mendenhall, C. D., Barrett, L. F., & Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24*(6), 947–956. https://doi.org/10.1177/0956797612464242

Yeung, M. K. (2022). A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: The specificity of deficits and the role of task characteristics. *Neuroscience and Biobehavioral Reviews, 133*, Article 104518.

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology, 31*(17), R1026–R1032. https://doi.org/10.1016/j.cub.2021.07.044

3. Publication II: A Virtual Reality Based System for the Screening and Classification of Autism

A Virtual Reality Based System for the Screening and Classification of Autism

Marta Robles*, Negar Namdarian*, Julia Otto*, Evelyn Wassiljew,
Nassir Navab *Fellow, IEEE*, Christine M. Falter-Wagner, Daniel Roth *Member, IEEE*

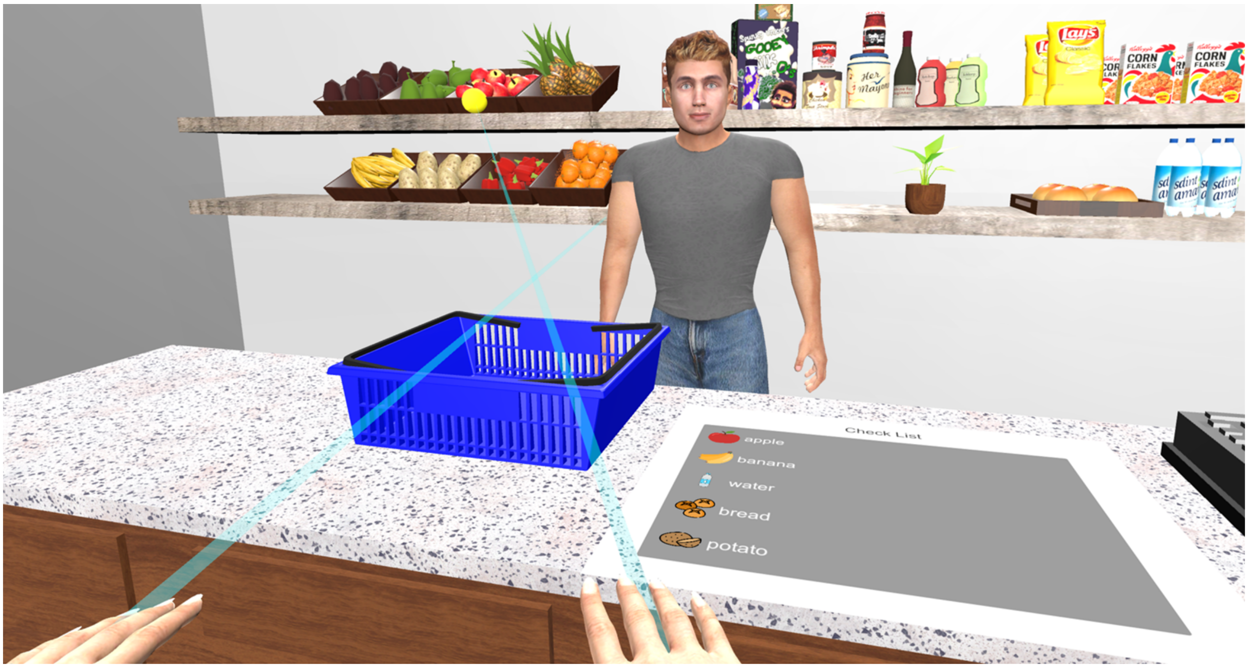

Fig. 1. **Example of the trial phase of the simulation.** A participant, embodied by an avatar, picks an item on the shopping list. After providing the item, the embodied agent representing the shop salesperson describes details and background information on the item, initiating a social interaction, resulting nonverbal responses from the participant. We record these nonverbal behaviors (e.g., head motion, eye gaze, gaze focus) and are able to classify autistic responses with high accuracy.

Abstract—Autism – also known as Autism Spectrum Disorders or Autism Spectrum Conditions – is a neurodevelopmental condition characterized by repetitive behaviours and differences in communication and social interaction. As a consequence, many autistic individuals may struggle in everyday life, which sometimes manifests in depression, unemployment, or addiction. One crucial problem in patient support and treatment is the long waiting time to diagnosis, which was approximated to seven months on average. Yet, the earlier an intervention can take place the better the patient can be supported, which was identified as a crucial factor.
We propose a system to support the screening of Autism Spectrum Disorders based on a virtual reality social interaction, namely a shopping experience, with an embodied agent. During this everyday interaction, behavioral responses are tracked and recorded. We analyze this behavior with machine learning approaches to classify participants from an autistic participant sample in comparison to a typically developed individuals control sample with high accuracy, demonstrating the feasibility of the approach. We believe that such tools can strongly impact the way mental disorders are assessed and may help to further find objective criteria and categorization.

Index Terms—Virtual reality, autism, machine learning, agents, embodiment, diagnosis

1 Introduction

Autism (ICD-11 6A02) is an entity used to define a set of persistent symptoms throughout the life cycle, characterised by (1) differences in communication and reciprocal social interaction and (2) the presence of repetitive behaviors and restricted interests [4, 52]. Concretely, the nonverbal communication skills of autistic individuals are particularly different compared to those of typically developed (TD) individuals. Nonverbal communication includes aspects such as initiating, maintaining or modulating gaze during a social interaction, modulating one's tone of voice when speaking or using gestures to accompany speech, among other features. The causes of this condition remain unclear (for a recent review, see [28]) but eye gaze patterns in autism have been pointed out as possible biomarkers of this condition [11, 16, 18]. More

- *Authors contributed equally to the project.
- Marta Robles, Christine Falter-Wagner, and Evelyn Wassiljew are with the Department of Psychiatry and Psychotherapy, Medical Faculty, LMU Munich. Marta Robles is also with Department of Clinical and Health Psychology, Autonomus University of Barcelona (UAB).
- Nassir Navab, Negar Namdarian, and Julia Otto are with TU Munich.
- Daniel Roth is with Friedrich-Alexander-Universität Erlangen-Nürnberg, Department Artificial Intelligence in Biomedical Engineering, and is the corresponding author (d.roth@fau.de).

Manuscript received 6 Sept. 2021; revised 3 Dec. 2021; accepted 7 Jan. 2022. Date of publication 16 Feb. 2022; date of current version 29 Mar. 2022. Digital Object Identifier no. 10.1109/TVCG.2022.3150489

concretely, eye-tracking has been used to investigate the gaze patterns of autistic individuals (for a systematic review and meta-analysis see [54]). The results of the mentioned review reveal that children with autism spectrum disorders (ASD have significantly reduced gaze fixation to the eye region of faces, when compared to TD individuals. Autism is one of the most prominent and widely discussed human conditions [31]. Despite of studies showing that diagnosis can be reliably established from the age of 2 years [43], many people with autism remain without a diagnosis, unrecognized, until adulthood [37]. Moreover, the average time to diagnose autism in adults has been estimated to 7 months [55], leading to long waiting lists and time. Further, the reliability of the common assessments for the diagnosis of autism, such as the Autism Diagnostic Observation Schedule (ADOS-2; [42], seems to be lower in adulthood [17, 44]. Not surprisingly, the diagnostic of autism in adulthood is one of the ten priority areas for autism research as published by Autistica [14]. Thus, there is a need for an objective measure to provide with a reliable and time economic diagnostics of ASD in adulthood.

1.1 Contribution

We present a system combining an agent-induced social virtual reality (VR) interaction with nonverbal behavior recording and pattern classification. We believe our approach could aid the diagnosis of autism and argue that it could be adapted in the future to also assist the screening of other social and communicative conditions or disorders. Our results are promising regarding the successful classification of autism based on a machine learning model trained and tested with the recorded data.

2 Background and Related Work

2.1 Virtual Environments and Autism Research

The use of VR technologies in autism research and therapy has grown in recent years, due to the strong level of experimental control. To date, VR has often been applied to interventions for children and adolescents on the autistic spectrum. For example, in the context of social communication, interaction and skill training [47, 57, 72], the training of emotion recognition, facial expression, as well as body gestures [22], phobia interventions [45, 46], the practice of fine motor skills [71] and driving exercises [66], (see [9] for a recent review). In populations with neurodegenerative diseases [53] or attention deficit hyperactivity disorder (ADHD) [2, 3], VR technologies have also been used as a tool to aid the evaluation processes to diagnose these conditions. Nowadays, the diagnosis is usually made in the clinic based on visible clinical signs and symptoms, and patients often have to wait for years for a correct diagnosis [53]. In the case of autism, this can be even more difficult, as the clinical heterogeneity of this condition is well known [10, 48]. This frequently leads to long evaluation processes including patients having to visit different experts, misdiagnosis and improper treatment [5]. Of course, these aspects have an impact on the patient's mental health [38].

In terms of ASD screening, few studies focused on the use of modern or novel technologies for ASD assessment. Koirala and colleagues [36] were the first to explore sensory abnormalities in ASD children with VR technology, whereas the automatic detection of ASD individuals revealed preliminary significant results in their study. VR in particular has shown an enormous contribution in clinical populations, in which eye-tracking on its own was only made possible to a limited extent [13]. Compared to regular VR devices, as well as eye-tracking technologies, immersive VR provides ecological validity in controlled environments by enabling a natural experience for the participants and therefore more reliable data collection [50].

2.2 Machine Learning-Based Autism Investigations

In medical context, machine learning (ML) has successfully been applied to objectively diagnose many different kinds of diseases and disorders, skin cancer [20] and heart diseases [56] being only two examples. In recent years, there has been a growing development of computer-aided investigations of ASD through ML on the basis of static images (e.g., [21, 32, 40, 41, 70]). Further, interpersonal synchronicity [23] has been investigated and classified with real-world motion data using motion energy analysis with a classification accuracy of 75.9%. Drimalla and colleagues [18] demonstated that a classification of facial behavior recorded from a video-based simulated dialogue study led to 73% accuracy in detecting ASD. Similarly, Yaneva and colleagues [69] could also detect autism automatically with around 74% accuracy. However, their approaches were based on a simulated interaction with a pre-recorded video [18] and in web page searches [69], which may not be fully capable to account for the full dynamics of social interactions.

In this regard, previous works presented potential methods and concepts to assess ASD using virtual characters [24, 61]. Further, specific study platforms for a potential behavior investigation have been developed [23, 59, 63]. In a recent study, Roth and colleagues [58] could automatically classify autistic individuals from a sample of ASD individuals and TD participants with up to 92.9% accuracy using a neural network trained from nonverbal cues from eye gaze and head movements recorded from avatar-mediated, dyadic social interactions in a desktop environment. While Georgescu and colleagues used motion analysis technologies to classify behaviors of real interactions [23], Roth and colleagues tracked the behavior of interactions that happened between two real people that were remotely tracked and represented to each other as avatars on desktop screens [58]. In contrast to these works, our goal was to implement a single user VR scenario that could allow to collect nonverbal data automatically, replicable, with high validity and experimental control.

3 Approach and Implementation

3.1 Virtual Environments and Scenarios

In order to create a virtual environment suitable for ASD assessment in VR, we identified different social settings that could be used for standardized social interactions. We used Autodesk 3ds Max[1], Blender[2], and the native tools in Unity 3D[3] to create our virtual environment. Some of the 3D models were acquired from Sketchfab[4].

Following a design discussion with clinical partners, we decided to implement a everyday life situation and standardized tasks in a virtual supermarket. For individuals with autism, shopping is a challenging daily living skill. When faced with an unfamiliar environment, such as at the supermarket, it was shown that diagnosed individuals show altered behaviors and affect [1]. Therefore, we anticipated different behavior from individuals with ASD while engaging in this simulation. Fig. 3 shows the final version of the shop used in our user study.

A social setting was considered essential for eliciting authentic nonverbal responses from participants during the simulation. Therefore, we created a virtual agent to act as the social partner in the role of the shop seller. As part of the simulation, the participants were instructed to purchase items shown on a shopping list (see Fig. 6). Purchasing the item required asking the seller to deliver it by ray-cast pointing and selection using the HTC Vive Controller. In the case of a correct selection, the shop seller agent would pick up the selected product and put it into the shopping basket. Following this action, the agent was designed to initialize a social interaction by narrating a short story or facts about the sold item accompanied by nonverbal behaviors. The narratives were co-designed with the clinical partners to match the right level between factual information and social engagement.

Taking into account the fact that different users may have varying levels of experience with VR, we created an introductory level within our simulation (see Fig. 2). Participants were presented with series of tutorials which covered the interaction with virtual objects, and how to adjust the volume of their headphones, etc. After finishing the tutorials, participants could see a start button on display, which could take them to the next phase of the simulation (see Fig. 2). Several studies have utilized virtual mirrors in order to increase the perception of embodiment toward a user's avatar [25, 60]. As a second part of the tutorial phase, the participants were therefore exposed to a virtual mirror to increase their awareness toward their presence and avatar within the simulation and to understand that their body behaviors are replicated and thus foster natural responses.

[1] Autodesk, 2020, San Rafael, USA. autodesk.com/products/3ds-max/
[2] Blender foundation, 2020. blender.org
[3] Unity Technologies, 2020, San Francisco, USA. unity.com
[4] Sketchfab, 2020, New York, USA. sketchfab.com

Fig. 2. **Tutorial Phase of the experiment.** Left: Tutorial environment. Center: Controller instructions. Right: Exposure to the virtual mirror.

Fig. 3. **Virtual shop environment.** The final design of virtual supermarket and surrounding scenario used in the study.

Fig. 4. **Virtual characters.**The virtual characters used for female participants (left), male participants (center) and the shop seller agent (right) in the study.

3.2 Avatar and Agent

Previous work reported that the perception of and interaction with virtual characters can be similar to a face-to-face interaction [15]. Therefore, it was argued that virtual characters, used as avatars (i.e., controlled by human behavior [6]) and agents (i.e., controlled by computer algorithms [6]) may act as a method to investigate social interactions in experimentally controlled fashion [24].

We used male and female virtual characters to represent the participants in the simulation accordingly. The avatar's height could be adjusted based on the participant's height. An HTC VIVE Pro head mounted display (HMD) with HTC VIVE controllers and HTC VIVE trackers (see Fig.6) in combination with SteamVR and the Unity Vive Input Utility allowed for the rendering as well as inverse kinematic tracking [62] in order to replicate the user behaviors to the avatar. Therefore, the participant's avatar's body movements corresponded to the participant's movements.

Similarly, we used a male virtual character as representation for the embodied agent (see Fig. 4). All characters were created using Autodesk Character Generator[5]. For the agent's motion and behavior, we integrated an animation state machine to drive the agent's action according to the current simulation status. We used both, Unity's internal animation system with customized keyframe animations (e.g., grabbing the products from the shelves, controlled arm rotations) as well as third party animation clips from the Unity Asset Store[6] and Mixamo[7] in order to construct all varieties of the agent's behavior. Each animation served as a state within the state machine, and various events could trigger the transition between these states. We took into account several factors such as velocity, and range of movement, to ensure smooth transitions.

To realize a more realistic gaze interaction, the virtual agent (seller) was capable of maintaining eye contact with the participant. Gaze shifting toward the participant involves eyes, head, and upper body movements, and eye, eyelid, blinking, as well as head animations and realistic lip-synced mouth movements were realized using a natural motion plugin (SALSA LipSync Suite Version 2.5.0.).

For the agents verbal discourse, a natural human voice was recorded for the verbal interaction. In order to have a voice that conveyed realistic emotions, the performer adapted his speech in accordance with the narrative.

3.3 Scenario and Logic

The agent's actions are triggered when the user selects an item with HTC VIVE handheld controller. Fig. 5 shows the flowchart of the avatar's actions in relation to the task procedure and status. If the participant selects an item that is not on the shopping list, the agent will ask them to try again. In the case that participants select an item on the list, the agent would bring the products to them. Following this, the agent will narrate a short story or fact about the sold item for the purpose of initiating social interaction. The participant could not select any other item while the agent is narrating the story. Once the final product is delivered, the agent will request payment. For payment, the participant has to drag a pack of cash visible on the counter towards the cash register and drop it there. After successful payment, the game will end with the agent saying goodbye to the participant.

3.4 Data Acquisition and Logging

Our system is designed to collect and log data during the social setting of the shopping scenario, which is the period in which the seller agent narrates a story for the participant. Since our research focused mainly on gaze, head, and body motion comparison between individuals with ASD and TD controls, we collect the body movements and eye gaze data of the participants. The body movement data include the position and rotation of head and hands. The gaze data was collected logging the gaze focus point along each axis (i.e., the gaze focus point in the 3D world) as well as the dwell time the participants focused on dynamic AOIs, virtually attached to landmarks of the agents face and hands, see Fig. 7. Once the raycast hits one of the colliders, the AOI, which was looked at by the participant, is detected. To prevent repetition, once a collision with an inner AOI (such as the eyes) is logged, the larger

[5]https://charactergenerator.autodesk.com/
[6]www.assetstore.unity.com
[7]www.mixamo.com

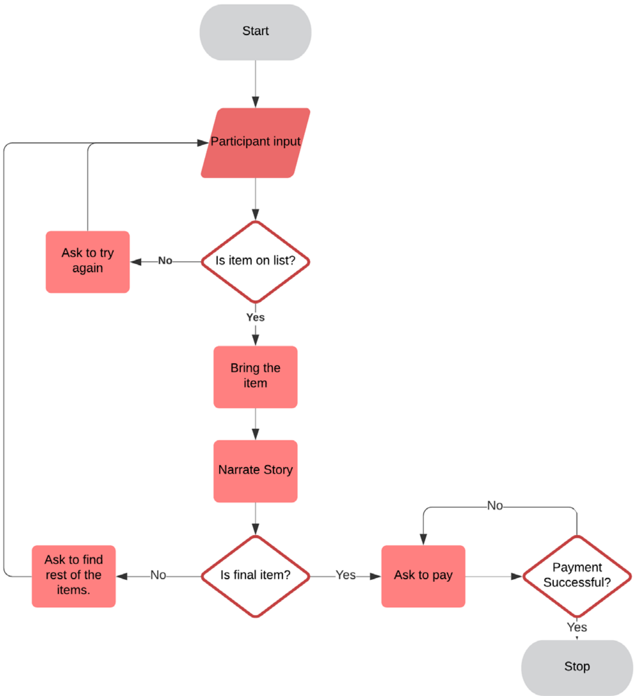

Fig. 5. **Scenario procedure.** The flowchart describes the simulation logic and user actions/agent reactions.

Fig. 6. **Embodiment method.** HTC VIVE Pro Eye and HTC VIVE trackers that allow for the embodiment of the user in the simulation by using inverse kinematics and body pose solving.

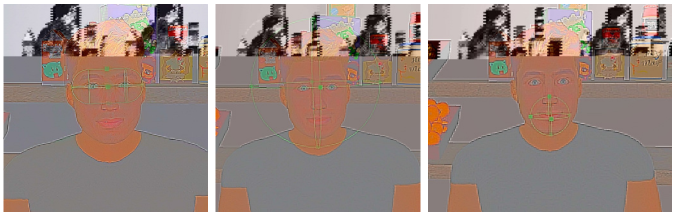

Fig. 7. **Dynamic areas of interest.** Dynamic head, eyes, and mouth AOIs assigned to the agent in order to acquire the focus dwell times.

area (such as the whole face) is ignored. During the virtual reality scenario a total of 27 features, including hand, head, and gaze behavior, were recorded using the HTC VIVE Pro Eye VR System and the Tobii XR SDK, respectively. We exported the acquired data into a CSV file format for analysis and further processing with Python and using Scikit-learn and tensorflow as machine learning libraries.

4 Evaluation

Based on previous research findings regarding the gaze pattern as a valid predictor for ASD [54] and based on our previous work from [58], it was suggested that *H1: there is a difference in mean fixation times on the eyes area, mouth area and the background during social interaction with a virtual character between ASD and TD individuals.* Our main research question was, in consequence, *RQ1: Can we classify ASD on the basis of the expressed nonverbal behavior (gaze, voice, head motion) acquired through an patient-agent system?*

4.1 Design and Task

For the data acquisition, we employed a between-group design and tested individuals with ASD vs. TD control participants. We used the virtual environment to simulate a social situation in which the nonverbal data from participants can be acquired and recorded to provide reliable data set for machine learning algorithms. As part of the simulation, the participants were instructed to purchase items shown on a list in the virtual environment representing a supermarket by pointing to and selecting the product in the shop using HTC VIVE handheld controllers via the controller's trigger button. This was available with both controllers accounting for different handedness of the participants. In the case that participants made the correct selection, the agent would bring the products to them and narrate about the sold item. An example would read as follows:

"Oh, these bananas are great. Did you know that bananas are rich in minerals such as magnesium, potassium and folic acid? Also, bananas are rich in vitamins B and C. And I tell you something: These bananas come from Ecuador. Ecuador has the perfect climate for its cultivation and exports every year more than six million bananas to all countries of the world."

During this time, the integrated eye-tracking system would gather data of the participant's eye gaze, and the transformation data from the head mounted display and controllers was recorded. The participants had to buy five items, which were presented in the same order for all participants. The narratives about the items were neutral facts and had a duration of about 90 seconds on average. Each data file collected during the scenario would then contain approximately 1850 measurement rows containing position, rotation and gaze values as well as the corresponding area of interest at the given moment during the simulation.

Taking into account the fact that different users might have varying levels of experience with VR, we created an introductory level within our simulation. To accomplish this, we designed a virtual space similar to an entrance to a market, to provide participants with the opportunity to become familiar with virtual technologies and understand how controllers work. An additional goal of the introductory level was to increase participants' sense of embodiment. During the simulation, participants were able to control an avatar representation of themselves. A study by Slater and Steed [64] showed that participants who interacted with virtual objects via a virtual body had a higher sense of presence in comparison to those using a traditional interface (like pressing a button) as a means of interaction. Virtual embodiment can lead to psychological effects such as increased social presence in users that control the avatar [60, 65]. In an attempt to evoke virtual embodiment, we considered the virtual mirror metaphor in our design. In this metaphor, users can see a simulated mirror reflection of their avatar. Several studies have used and tested virtual mirrors. A study by González-Franco and colleagues [25] concluded that seeing the avatar reflection of oneself in a virtual mirror, while the movements are synchronous with the user, would result in a higher subjective sense of embodiment. Assuming that a greater perception embodiment would also result in more natural behavioral responses, we implemented a mirror in the introductory part of the virtual simulation.

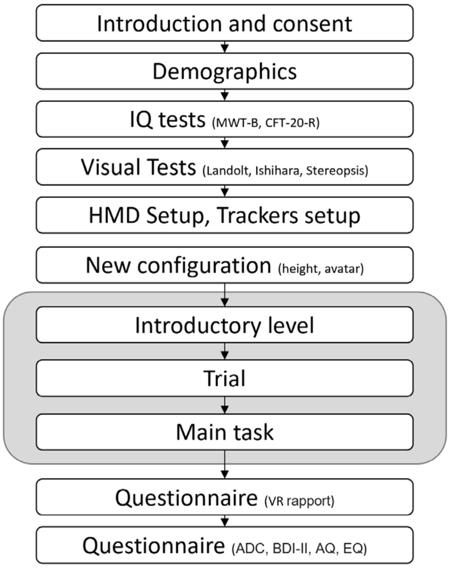

Fig. 8. **E**xperimental procedure.

4.2 Procedure

The experimental session can be roughly divided into three phases. First, the participants answered a batch of psychological and optical questionnaires and tests, to confirm the inclusion criteria of the participants (including a Landolt C-test, Ishihara Color test, and Steropsis test). Then, an eye calibration was performed and the VR simulation was conducted. We informed participants about potential cybersickness and that asked them to immediately notify investigators. The VR task lasted about 20-30 minutes. After the participants completed the task, they completed further questionnaires related to the VR task as well as psychological questionnaires. The session lasted for approximately 90 to 120 minutes. The participants were each compensated financially for taking part in the study. All participants gave written informed consent before study participation. The study was approved by the ethics committee of the Ludwig-Maximilians University Munich Hospital, in agreement with the Declaration of Helsinki [68]. The full procedure is depicted in Fig. 8.

4.3 Measures

The German versions of the following questionnaires where administrated: a demographic questionnaire, to collect biographical data and previous VR experience; the Multiple Choice Vocabulary Intelligence Test (verbal intelligence; MWT-B; [39]) and the Basic Intelligence Scale Scale 2 - Revision (non-verbal intelligence; CFT-20-R; [67]) to corroborate inclusion criteria of the participants' IQ; the Landolt C test, the Stereo Optical test [51] and the Ishihara colour-blindness test [12] for ocular and colour deficiencies. In phase three, after the VR task, participants were asked to answer a rapport questionnaire to assess the quality of the interaction with the avatar [35]; the autism-spectrum-quotient (measuring extent of autistic traits; AQ, [8]) and the empathy quotient (to assess empathy; EQ, [7]). In addition, we measured motor difficulties with the Adult Dyspraxia Checklist (ADC) [33] and depressive symptoms with the Beck Depression Inventary (BDI-II) [27]. We will not go into detail of this reporting due to the fact that the underlying research questions are not in the focus of the present study.

4.4 Participants

A total of 28 participants took part in the study. Twenty TD participants were recruited via social networks and acquaintances. Eight individuals with a clinically confirmed diagnosis of ASD were recruited through the specialised autism outpatient clinic of the University Hospital of Munich. We excluded participants who stated or reported being very tired since their gaze paths are likely to be altered, as well as participants who were distracted or did not follow the task instructions, leading to 6 TD control and 2 ASD participants data sets being excluded from the dataset. In addition, one male TD participant had to be excluded as technical issues led to data distortion. Therefore, the final sample was composed by 13 TD (9 female, 4 male, age $M = 23.31$, $SD = 2.39$) and 6 ASD (3 female, 3 male, age $M = 28.83$, $SD = 8.98$) participants ($N = 19$). Descriptive statistics can be found in Table 1.

	ASD	Matched	Random
Age	28.8 (8.9)	23.16 (2.0)	23.5 (2.5)
Verbal IQ	110.0 (5.0)	108.0 (14.2)	104.16 (10.6)
Non-verbal IQ	110.66 (16.45)	114.1 (14.5)	114.1 (13.7)

Table 1. Descriptive statistics: M (SD) of participants data of the matched control and the random control data sets compared to the ASD set.

	t	df	p	$Cohen's\ d$	95% CI Lower	95% CI Upper
Gender	0.00	10.00	1.00	0.00	-1.132	1.132
Age	1.508	10.00	0.163	0.871	-0.341	2.044
Verbal IQ	0.324	10.00	0.753	0.187	-0.952	1.317
Non-verbal IQ	-0.390	10.00	0.705	-0.225	-1.356	0.916

Table 2. T-test results (matched set).

5 Results

5.1 Analysis Strategy

To analyse collected data for a balanced group comparison that is better applicable to machine learning classification approaches, six TD participants (3 female, 3 male, age $M = 23.16$, $SD = 2.0$) were case-wise matched to ASD participants based on age, IQ, and gender, see Table 1. T-test results showed that participants could be matched on the basis of gender and IQ scores, see Table 2). For the age, Levene's test was significant ($p < .05$) suggesting a violation of the equal variance assumption.

Additionally, a separate analysis based on matching the 6 ASD participants with 6 randomly chosen TD participants was conducted (see Table 1). Table 3 shows that the gender and IQ scores in random selection matched between the two groups, but the age did not. In the following we report both, the ASD vs. matched TD control comparison as well as ASD vs. random TD control comparison.

5.2 Descriptive Analysis and Comparisons

As expected, both groups showed differences in the AQ and EQ tests: the autistic individuals and the TD individuals in the matched data set comparison did defer in autistic traits (AQ: $t(10) = 2.09$, $p = .064$, $Cohen's\ d = 1.204$, 95% CI from 0.066 to 2.425) but not to a significant level, however did significantly differ in empathy skills (EQ: $t(10) = -2.48$, $p = .033$, $d = -1.429$, 95% CI from -2.692 to -0.113). In the random data set comparison, the groups deferred significantly in both constructs (AQ: $t(10) = 2.70$, $p = .022$, $d = 1.557$, 95% CI from 0.212 to 2.846; EQ: $t(10) = -2.43$ $p = .035$, $d = -1.404$, 95% CI from -2.661 to -0.093). Analysis of nonverbal behavior data collected during the monologues of the VR simulation was compared in between the two groups for both the matched and the random data set. As expected, the differences in gaze behavior were significant regarding the average dwell time on the eyes (matched group

	t	df	p	$Cohen's\ d$	95% CI Lower	95% CI Upper
Gender	0.542	10.00	0.599	0.313	-0.834	1.445
Age	1.398	10.00	0.192	0.807	-0.395	1.973
Verbal IQ	1.210	10.00	0.254	0.699	-0.489	1.854
Non-verbal IQ	-0.400	10.00	0.698	-0.231	-1.361	0.911

Table 3. T-test results (random set).

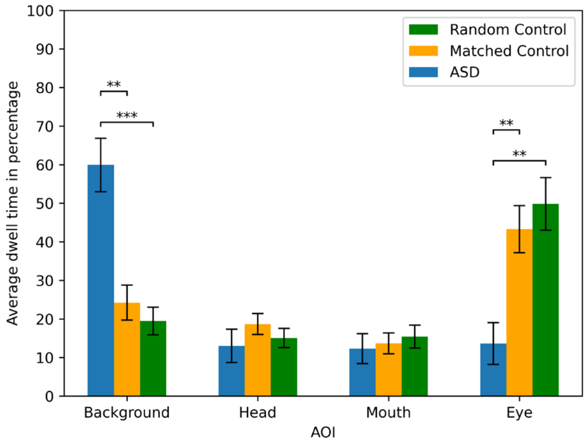

Fig. 9. **AOI dwell times.** The average dwell time on AOIs of the matched and the random set of participants. Note: Graphs denote $M \pm SE$. Asterisks denote significant differences resulting from Student independent t-tests. * indicate $p - value < .05$, ** indicate $p - value < .01$, *** indicate $p - value < .001$

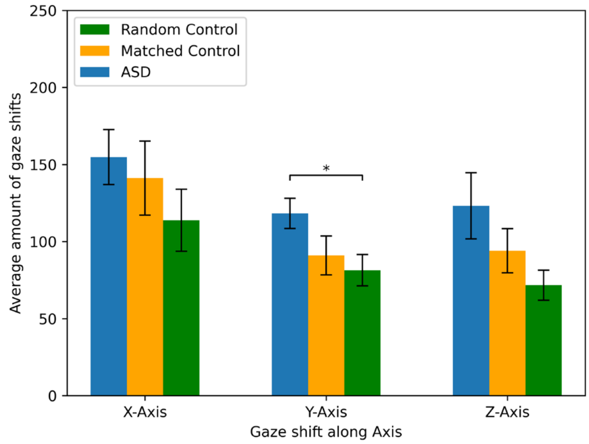

Fig. 10. **Gaze shifts.** The averages of the absolutes of the sum of gaze shifts of the focus point in 3D space *relative* to the head orientation (i.e., local translation) of the matched and the random set of participants. Note: Graphs denote $M \pm SE$. Values represent vector values not axis rotations. Z equals to the view direction, Y is the up axis. Asterisks denote significant differences resulting from Student independent t-tests. * indicate $p - value < .05$, ** indicate $p - value < .01$, *** indicate $p - value < .001$

comparison: $t(10) = -3.64$, $p = .005$, $d = -2.103$, 95% CI from -3.523 to -0.620; random group comparison: $t(10) = -4.17$, $p = .002$, $d = -2.405$, 95% CI from -3.909 to -0.837) and the background (matched group comparison: $t(10) = 4.32$, $p = .002$, $d = 2.494$, 95% CI from 0.900 to 4.025; random group comparison: $t(10) = 5.19$, $p < .001$, $d = 2.997$, 95% CI from 1.246 to 4.685), see Fig. 9. Results suggest a gaze shift towards the background for the ASD participants, whereas their focus on the eye region is reduced. However, we did not find the expected longer focus on the mouth region ($p > .05$). The differences in gaze shifts of the focus point in head-relative 3D space were not significant regarding X-axis (matched group comparison: $t(10) = 0.46$, $p = .657$, $d = 0.265$, 95% CI from -0.879 to 1.395; random group comparison: $t(10) = 1.53$, $p = .157$, $d = 0.883$, 95% CI from -0.331 to 2.057), Y-axis for the matched group comparison ($t(10) = 1.71$, $p = .118$, $d = 0.987$, 95% CI from -0.244 to 2.175) but significant for the random group comparison ($t(10) = 2.612$, $p = .026$, $d = 1.508$, 95% CI from 0.174 to 2.786), and again, non-significant for Z-axis (matched group comparison: $t(10) = 1.13$, $p = .286$, $d = 0.650$, 95% CI from -0.531 to 1.801; random group comparison: $t(10) = 2.18$, $p = .054$, $d = 1.259$, 95% CI from -0.022 to 2.490), see Fig. 10. Finally, the differences in head rotation were not significant regarding X-axis for the matched group comparison ($t(10) = 1.08$, $p = .305$, $d = 0.625$, 95% CI from -0.553 to 1.774), but significant for the random group comparison ($t(10) = 2.32$, $p = .043$, $d = 1.342$, 95% CI from 0.044 to 2.587) and non-significant for Y-axis (matched group comparison: $t(10) = 0.76$, $p = .468$, $d = 0.436$, 95% CI from -0.722 to 1.573; random group comparison: $t(10) = 1.84$, $p = .096$, d = 1.062, 95% CI from -0.182 to 2.260, and Z-axis (matched group comparison: $t(10) = 0.17$, $p = .871$, $d = 0.096$, 95% CI from -1.038 to 1.226; random group comparison: $t(10) = 0.69$, $p = .509$, $d = 0.396$, 95% CI from -0.758 to 1.531), see Fig. 11.

5.3 Preprocessing

Data files collected during the experiment were preprocessed. Foremost, we removed any invalid data due to tracking (system) errors. Invalid data for example arises when the tracker can not detect eye movement. The average amount of invalid data detected during the testing phase of the simulation and also the user study is less than 10% of the collected data frames per simulation (ASD: 6.7%, matched controls: 8.87%, random controls: 6.76%). We removed this data from the dataset. To validate approaches of previous works [23, 58], we transformed gaze vectors to present the gaze shift in local coordinate space and calculated gaze averages and further calculated averages for all features.

A similar approach was followed for the time series (i.e., individual monologue) based data set that were analyzed using a LSTM classification approach. We separated each monologue during the simulation and calculated the averages. An overview of the preprocessing is provided in Fig. 12. Most machine learning algorithms have difficulty handling largely varying scales of input features. Therefore, we scaled the data for all features in a limited range by min-max scaling, as it transforms all values to the range $[0, 1]$, which is the expected input for most neural network algorithms [26].

5.4 Classification

We used similar parameters for the logistic regression, support vector machine, and neural network than previous work. We further implemented an LSTM based on the data of the individual monologues. We chose a sigmoid activation function and binary cross entropy as loss function, and a stochastic gradient descent (SGD) as optimizer. The training set consisted of 80% of the available data while the test set evaluated for validation contains the remaining 20%. Both sets contained an equal amount of TD control and ASD sample data. The machine learning pipeline is depicted in Fig. 13.

5.4.1 Validation of Previous Findings

In order to quantify results of previous work and compare results of this thesis, previously implemented algorithms of a similar setting are tested on the new data [58]. In named study, autism is classified through application of three different types of machine learning algorithms, including logistic regression, a support vector machine and a neural network.

Each algorithm was evaluated applying 5-fold cross validation.

The logistic regression model was trained on all extracted features. Results from training the model on data collected during the user study, reveal an average accuracy of 80% ($SD = 0.4$), sensitivity of 80% ($SD = 0.4$) and specificity of 80% ($SD = 0.4$), with $C = 0.5$ and a maximum of 5000 iterations for the matched data set as well as the random data set.

Training the support vector machine on data from the user study achieves an average accuracy of 63.3% ($SD = 0.306$), sensitivity of 80% ($SD = 0.4$) and specificity of 60% ($SD = 0.49$) for the matched

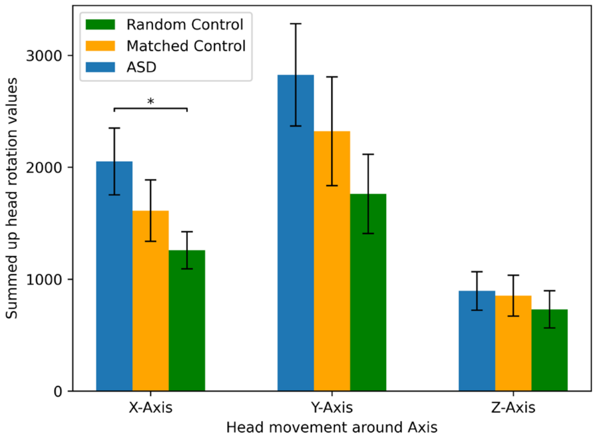

Fig. 11. **Head motion.** Sum of the total head rotation averages per group for each axis. Z equals to the view direction, Y is the up axis (i.e., captures the horizontal head rotation). Note: Graphs denote $M \pm SE$. Asterisks denote significant differences resulting from Student independent t-tests. * indicate $p-value < .05$, ** indicate $p-value < .01$, *** indicate $p-value < .001$

data set and an accuracy of 80% ($SD = 0.4$), sensitivity of 80% ($SD = 0.4$) and specificity of 80% ($SD = 0.4$) for the random data set. Parameters for the matched set were $\gamma = 0.001$, $C = 9.9$ and $\gamma = 0.009$, $C = 9.9$ for the random set.

An neural network consisting of one hidden layer achieved an accuracy of 76.7% ($SD = 0.29$), sensitivity of 80% ($SD = 0.40$) and specificity of 70% ($SD = 0.40$) for the matched data set with $hidden\_layer\_size = 5$. An accuracy of 93.30% ($SD = 0.13$), sensitivity of 80% ($SD = 0.40$) and specificity of 1.0 ($SD = 0.00$) for the random data set is achieved with $hidden\_layer\_size = 6$. On the other hand, an average accuracy of 86.70% ($SD = 0.16$), sensitivity of 70% ($SD = 0.40$) and specificity of 100% ($SD = 0.00$) is achieved with $hidden\_layer\_size = (6,21)$ for user study data matched set is achieved by an ANN consisting of two hidden layers and an accuracy of 93.30% ($SD = 0.13$), sensitivity of 80% ($SD = 0.40$), specificity of 100% ($SD = 0.00$) and $hidden\_layer\_size = (7,23)$ is achieved for the random set. Accuracy results are compared in Table 4

Testing previous algorithms on a data set not consisting of averages of the full conversation but by evaluating each conversation of the simulation separately reveals a more defined accuracy (see Table 5). The new data consists of a total of 60 sets instead of 12 as the simulation consists of five monologues.

5.4.2 Classification Using a LSTM

An LSTM network consisting of a single hidden layer was implemented. The best parameters for learning rate (0.1) and epochs (250) were chosen by test and evaluation. The LSTM achieved an accuracy of 100%, sensistivity of 83.0% and a specificity of 99.1% and 98.9% on all features, equally for the matched and random data set.

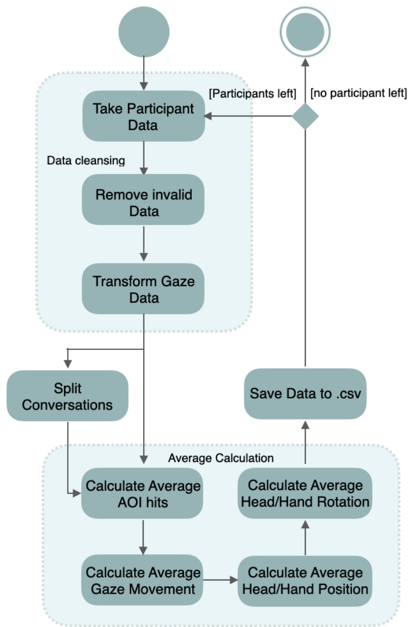

Fig. 12. **Preprocessing.** Overview of the data preprocessing including exclusion of invalid data and average calculation.

Feature Analysis In order to reduce computing time and increase accuracy, it is beneficial to look at different smaller feature combinations separately instead of evaluating the LSTM on all features at once. Some features might be stronger indicators for autism than others and some might not show a significant difference between the two groups. Those could in turn be left out of the calculation, saving computing time in the process. Accuracy is calculated for all possible two and three feature combinations of the 27 features of the collected data. The accuracy the algorithm can achieve is calculated for each single feature as well. In Table 6 and Table 7 some examples of feature combinations of two and three for the matched data set can be seen. Evaluation on the random data set shows similar results, see Table 8 and Table 9. Combinations of features that include one of the AOIs in general perform better than combinations of features only consisting of position or rotation features. However, there is no clear combination winner in the two evaluations taking into account two features. Evaluation of one single feature achieved the highest accuracy for Background and Eye (see Table 10) for the matched data set and for only Background for the random data set.

Multilayer LSTM A multilayer neural network is expected to improve accuracy compared to a single layer network for large amounts of data and may reduce over-fitting. As the amount of data in this study is limited, the outcome may not pose a significant difference. Including a second layer in the model resulted in an accuracy of 100%, a sensitivity of 82.9% and a specificity of 98.7% for classifying autism correctly on the matched and an accuracy of 100%, a sensitivity of 81.4% and a specificity of 93.1% for the random data set, with $learning\_rate = 0.1$ and $number\_of\_epochs = 250$ for both sets and thus even under-performed the single layer approach regarding the

	ASD vs. Matched TD	ASD vs. Random TD
Logistic Regr.	0.80 (0.40)	0.80 (0.40)
SVM	0.80 (0.31)	0.80 (0.40)
MLP 1 Layer	0.77 (0.29)	0.93 (0.13)
MLP 2 Layer	0.87 (0.17)	0.93 (0.13)

Table 4. Accuracy M (SD) for each approach based on the evaluation of the averages of the full data set.

	ASD vs. Matched TD	ASD vs. Random TD
Logistic Regr.	0.93 (0.10)	0.93 (0.10)
SVM	0.93 (0.10)	0.95 (0.10)
MLP 1 Layer	0.97 (0.04)	0.97 (0.04)
MLP 2 Layer	0.98 (0.03)	0.98 (0.03)

Table 5. Accuracy M (SD) for each approach based on the evaluation of the averages of each monologue.

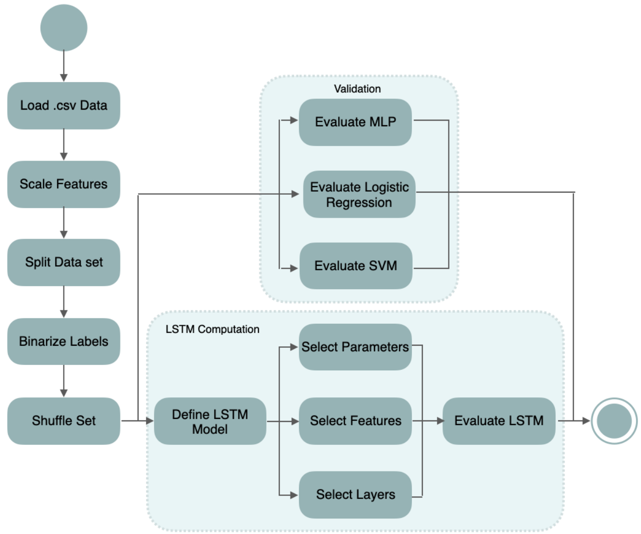

Fig. 13. **Classification procedure.** Representation of the machine learning pipeline depicting data preparation and evaluation of applied algorithms.

Features	Accuracy
(Background, Eye)	1.0
(Head, Gaze Vector Y)	1.0
(Mouth, Gaze Vector Y)	0.93
(Head Position Z, Left Hand Position Z)	0.87

Table 6. Accuracy of classification evaluating on possible two feature combinations for matched data set.

key performance indicators. However, we expect this approach to be more beneficial with larger data sets and the performance achieved can be considered sufficient for an assistive tool.

6 Discussion

In this paper, we present a VR-based system that could act as an assistive screening tool in objectively supporting the diagnosis of ASD and report on its evaluation. The simulation exposes the user to a social situation with an embodied virtual agent. We record resulting nonverbal behavior parameters of the user and use those to perform a classification judgement. With different approaches, we could reach accuracies up to 100% with our (limited) samples. While the sample size limits the generalizability of the approach due to the lack of a large general population representation, we interpret these results as very promising. Our procedure is not invasive, and we assume that, when put into practice, a screening test with this procedure would not take more than approximately 45 minutes till a result could be obtained.

Our subjacent research goal was to investigate whether and to what degree VR technology can support the ASD detection process by automatically distinguishing behavior and gaze characteristics between TD individuals and individuals with autism through VR and nonverbal pattern analysis/pattern classification techniques. In this regard, we are the first to show the successful implementation of a VR driven

Features	Accuracy
(Background, Mouth, Head Position X)	1.0
(Background, Eye, Gaze Vector Y)	1.0
(Left Hand, Mouth, Eye)	1.0
(Head Position Z, Gaze Vector Z, Right Hand Position X)	0.87

Table 7. Accuracy of classification evaluating on possible three feature combinations for matched data set.

Features	Accuracy
(Background, Head)	1.0
(Eye, Gaze Vector Y)	1.0
(Gaze Vector Y, Right Hand Rotation X)	0.93
(Head Position Z, Left Hand Rotation X)	0.87

Table 8. Accuracy of classification evaluating on possible two feature combinations for random data set.

Features	Accuracy
(Background, Left Hand, Mouth)	1.0
(Background, Head, Gaze Vector Y)	1.0
(Background, Head, Eye)	1.0
(Head Position Y, Left Hand Rotation Y, Left Hand Rotation Z)	0.87

Table 9. Accuracy of classification evaluating on possible three feature combinations for random data set.

screening tool and thus argue that found supporting results indicating a positive answer to our *RQ 1: Can we classify ASD on the basis of the expressed nonverbal behavior (gaze, voice, headmotion) acquired through an patient-agent system?*, although not without limitations.

Our results show strong differences in the mean fixation times in the *eye region* of the virtual agent and the *background region* of the grocery shop can be observed between participants with autism and TD individuals. Contrary to previous studies [49, 58] that showed differences in focus times of the mouth area, we did not observe a significant difference in this measure. One interpretation may be that in relation, the overall focus of ASD participants was mainly the background area such that the head and head area generally did not receive much attention at all, because sufficient context cues where available to avoid this area completely. We can thus only partially support *H1: there is a difference in mean fixation times on the eyes area, mouth area and the background during social interaction with a virtual character between ASD and TD individuals*. However, we implemented and confirmed other indicators that can contribute to the screening, such as the head position and the gaze vector in 3D space, that confirms the results found in a previous study [58]. To this end, non-verbal gaze behavior has been shown to be a notable factor in the recognition of autistic features and has been investigated for many years [19, 34]. Previous works investigated a desktop-based virtual environment prototype that classifies autism and thus revealed significant results in distinguishing between the two groups of adult subjects with respect to their gaze pattern, with high categorisation accuracies (up to 92.9%) [58]. Yet, previous works have been using either a a) a still picture based assessment or b) a dyadic interaction assessment, which may a) not account for the full and subtle dynamics of social interaction or b) require two participants or one participant and a therapist to be part of the procedure. With the present system and study we could substitute one participant posed by an avatar with an embodied virtual agent and minimise the setting requirements to a one person configuration, maintaining and confirming the previous works' performance and increase the level of accuracy and other key performance indicators. In comparison to previous work, we could also use a large percentage of the data collected without invalid data points, since the tracker is fixed to the head and integrated in the HMD, accounting for changes in head orientation that would lead to errors with regular desktop trackers.

A recent review suggest that receiving an autism diagnosis has a significant emotional impact on adults and that accessibility and processes are inconsistent [30]. Moreover, earlier diagnosis could prevent secondary mental health problems in this population [29]. We believe that the present study could assist this processes and improve non-objective and time consuming standard assessments. This study also contributes to the field of diagnosis research evidences, one of the ten priority areas for autism research [14]. We believe that a tool, such as ours, could not only be extended to include a broader population, but also to distinguish and identify other social and communicative disorders, such as Borderline or Schizophrenia, that manifest in differences and

Feature	Accuracy Matched TD controls	Accuracy Random TD controls
Background	1.00	1.00
Eye	1.00	0.73
Head	0.53	0.33
Mouth	0.33	0.33
Left Hand	0.33	0.33
Right Hand	0.33	0.33
Head Position X	0.33	0.33
Head Position Y	0.40	0.40
Head Position Z	0.86	0.33
Head Rotation X	0.33	0.33
Head Rotation Y	0.33	0.33
Head Rotation Z	0.33	0.33
Gaze Vector X	0.33	0.60
Gaze Vector Y	0.93	0.73
Gaze Vector Z	0.80	0.33
Left Hand Position X	0.46	0.46
Left Hand Position Y	0.33	0.33
Left Hand Position Z	0.93	0.66
Left Hand Rotation X	0.33	0.93
Left Hand Rotation Y	0.80	0.46
Left Hand Rotation Z	0.33	0.59
Right Hand Position X	0.33	0.33
Right Hand Position Y	0.60	0.33
Right Hand Position Z	0.33	0.33
Right Hand Rotation X	0.33	0.33
Right Hand Rotation Y	0.33	0.33
Right Hand Rotation Z	0.33	0.33

Table 10. Accuracy of classification evaluating on single feature for the matched TD control and the random TD control data set.

altered patterns of social and nonverbal behavior in everyday life.

In particular, we believe that including a tool such as the proposed one in a screening procedure could substantially i) reduce the waiting time by speeding up the initial procedure and pathway decision process, ii) reduce the patient and medical system costs, and iii) provide additional certainty and assurance to the therapist and is superior to other subjective questionnaire assessments.

6.1 Limitations

Of course, our study and results cannot be blindly generalized and are not without limitations. We recognise that the sample size of the present study is an obvious limitation, that keeps from generalizing the findings to be applicable to diverse screening populations. In addition, the limited sample size poses the risk of introducing overfitting in the neural network. Therefore, the results presented should be interpreted with caution and future research should be conducted to significantly increase the sample size to support and corroborate the results shown. Future recruitment may also consider the diversity of the population included, including patients with disorders that are related in the behavior manifestations. In addition, our TD sample was not screened and therefore we cannot exclude that participant with mental disorders are present in the sample (above or below the average percentage in the general population). However, all TD participants stated that there is no presence of any disorders.

Further, in the present study we only included adult participants with no intellectual impairments. Thereby, future research should both include a wider age range and individuals with different cognitive styles. In this regard however, it is necessary to change and redesign the 3D environment and simulation accordingly. Yet, our scenario and simulation principle offers dynamics to simplify the task or environment to a degree understandable for people with intellectual deficits and children in developing ages. Finally, our agent is not yet capable of initiating and maintaining bidirectional, i.e., interactive, social communication. Future work may include such interaction either by a screen based dialogue and answer selection or by simplified yes and no answers recognized with speech recognition. However, from the current data, we do see that differences can be shown even with the simplified interaction type we implemented.

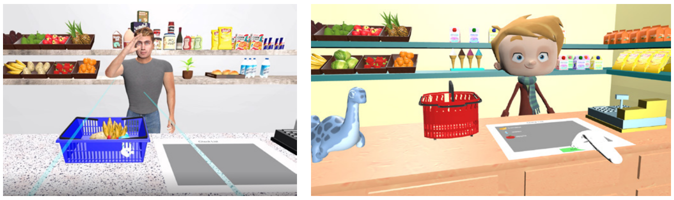

Fig. 14. **Future work.** Left: current environment. Right: Prototype of a potential adaptation of the scenario to younger samples.

6.2 Future Work

In future works, we aim to mitigate our limitations, to increase the sample size and include other disorders that manifest in behavioral differences. In that regard, our goal is to look for deep ML methods to be applied, such as by including the whole time series of the data collections. That would potentially enable another dimension, i.e. frame dependent measures, and may allow distinguishing between subtle differences in disorder types, or classify the severity within one neurodevelopmental disorder. Furthermore, we aim at adjusting our simulation characteristics for other populations, such as children, since it is especially important to have a diagnosis as soon as possible in order to establish interventions. Figure 14 presents a first impression on how style and simulation could be adapted for younger aged populations. The impact of the level of fidelty and naturalness of the communication behavior of the virtual agent leaves room for future endeavors. While our interaction was simplistic, it may be the target of future research to investigate more natural interactions, potentially allowing for a more realistic bidirectional communication and interaction. A more natural communication using a wizard of oz paradigm or a more intelligent agent may even improve the present results and system's performance. To this regard, future work may also consider to vary the number of agents and their proxemics, which was not subject of the current investigation. Rather, the interpersonal distance was chosen by an estimate of what would feel natural and physically sound. Finally, future research should also aim to clinically validate the presented screening tool, pursuing to classify the severity of the disorder.

7 Conclusion

Our proposed VR system for autism classifcation and the presented evaluation results showed that the system is capable of a nonverbal behavior pattern classification between autistic and typically developed individuals with a high accuracy, sensitivity, and specificity. Confirming previous studies, focus behavior as well as gaze movement were strong features. Our system could not only assist diagnostic procedures of autism but be extended and used for the assessment of other communicative and social disorders. We are convinced that our system could be successfully deployed as an assistive tool in the screening and diagnosis procedures to reduce waiting times and costs, as well as to provide an objective method of assessment.

Acknowledgments

The authors wish to thank Prof. Dr. Zhuanghua Shi from LMU Munich for his support during the data collection. We also thank Harry Unwin for creating the audio recordings for the agent.

References

[1] A. Adjorlu, E. R. Hoeg, L. Mangano, and S. Serafin. Daily Living Skills Training in Virtual Reality to Help Children with Autism Spectrum Disorder in a Real Shopping Scenario. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 294–302. IEEE, Nantes, France, Oct. 2017. doi: 10.1109/ISMAR-Adjunct.2017.93

[2] D. Areces, J. Dockrell, T. García, P. González-Castro, and C. Rodríguez. Analysis of cognitive and attentional profiles in children with and without ADHD using an innovative virtual reality tool. *PloS one*, 13(8):e0201039, 2018. Publisher: Public Library of Science San Francisco, CA USA.

[3] D. Areces, C. Rodríguez, T. García, M. Cueli, and P. González-Castro. Efficacy of a continuous performance test based on virtual reality in the

diagnosis of ADHD and its clinical presentations. *Journal of Attention Disorders*, 22(11):1081–1091, 2018. Publisher: Sage Publications Sage CA: Los Angeles, CA.

[4] A. P. Association and A. P. Association, eds. *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Association, Washington, D.C, 5th ed ed., 2013.

[5] S. K. Au-Yeung, L. Bradley, A. E. Robertson, R. Shaw, S. Baron-Cohen, and S. Cassidy. Experience of mental health diagnosis and perceived misdiagnosis in autistic, possibly autistic and non-autistic adults. *Autism*, 23(6):1508–1518, Aug. 2019. Publisher: SAGE Publications Ltd. doi: 10.1177/1362361318818167

[6] J. Bailenson and J. Blascovich. Avatars.

[7] S. Baron-Cohen and S. Wheelwright. The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, 34(2):163–175, Apr. 2004. doi: 10.1023/B:JADD.0000022607.19833.00

[8] S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin, and E. Clubley. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, 31(1):5–17, Feb. 2001. doi: 10.1023/A:1005653411471

[9] C. Berenguer, I. Baixauli, S. Gómez, M. d. E. P. Andrés, and S. De Stasio. Exploring the Impact of Augmented Reality in Children and Adolescents with Autism Spectrum Disorder: A Systematic Review. *International Journal of Environmental Research and Public Health*, 17(17):6143, Jan. 2020. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute. doi: 10.3390/ijerph17176143

[10] H. Bruining, L. d. Sonneville, H. Swaab, M. d. Jonge, M. Kas, H. v. Engeland, and J. Vorstman. Dissecting the Clinical Heterogeneity of Autism Spectrum Disorders through Defined Genotypes. *PLOS ONE*, 5(5):e10887, May 2010. Publisher: Public Library of Science. doi: 10.1371/journal.pone.0010887

[11] R. Camero, V. Martínez, and C. Gallego. Gaze Following and Pupil Dilation as Early Diagnostic Markers of Autism in Toddlers. *Children*, 8(2):113, Feb. 2021. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute. doi: 10.3390/children8020113

[12] J. H. Clark. The Ishihara Test for Color Blindness. *American Journal of Physiological Optics*, 5:269–276, 1924.

[13] V. Clay, P. König, and S. U. König. Eye tracking in virtual reality. *Journal of Eye Movement Research*, 12(1), Apr. 2019. Number: 1. doi: 10.16910/jemr.12.1.3

[14] J. Cusack and R. Sterry. Your questions: shaping future autism research. Technical report, Autistica, 2016.

[15] A. de Borst and B. Gelder. Is it the real deal? Perception of virtual characters versus humans: An affective cognitive neuroscience perspective. *Frontiers in Psychology*, 6, May 2015. doi: 10.3389/fpsyg.2015.00576

[16] M. Del Valle Rubido, J. T. McCracken, E. Hollander, F. Shic, J. Noeldeke, L. Boak, O. Khwaja, S. Sadikhov, P. Fontoura, and D. Umbricht. In search of biomarkers for autism spectrum disorder. *Autism research*, 11(11):1567–1579, 2018. Publisher: Wiley Online Library.

[17] A. der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF). S3-Leitlinie Autismus-Spektrum-Störungen im Kindes-, Jugend- und Erwachsenenalter Teil 1: Diagnostik. Technical report, DGKJP, DGPPN, 2015.

[18] H. Drimalla, T. Scheffer, N. Landwehr, I. Baskow, S. Roepke, B. Behnia, and I. Dziobek. Towards the automatic detection of social biomarkers in autism spectrum disorder: introducing the simulated interaction task (SIT). *npj Digital Medicine*, 3(1):1–10, Feb. 2020. doi: 10.1038/s41746-020-0227-5

[19] M. Elsabbagh, A. Volein, G. Csibra, K. Holmboe, H. Garwood, L. Tucker, S. Krljes, S. Baron-Cohen, P. Bolton, T. Charman, G. Baird, and M. H. Johnson. Neural correlates of eye gaze processing in the infant broader autism phenotype. *Biological Psychiatry*, 65(1):31–38, Jan. 2009. doi: 10.1016/j.biopsych.2008.09.034

[20] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, Feb. 2017. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7639 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Diagnosis;Machine learning;Skin cancer Subject_term_id: diagnosis;machine-learning;skin-cancer. doi: 10.1038/nature21056

[21] D. Fabiano, S. Canavan, H. Agazzi, S. Hinduja, and D. Goldgof. Gaze-based classification of autism spectrum disorder. *Pattern Recognition Letters*, 135:204–212, 2020. Publisher: Elsevier.

[22] S. Fridenson-Hayo, S. Berggren, A. Lassalle, S. Tal, D. Pigat, N. Meir-Goren, H. O'Reilly, S. Ben-Zur, S. Bölte, S. Baron-Cohen, and O. Golan. 'Emotiplay': a serious game for learning about emotions in children with autism: results of a cross-cultural evaluation. *European Child & Adolescent Psychiatry*, 26(8):979–992, Aug. 2017. doi: 10.1007/s00787-017-0968-0

[23] A. L. Georgescu, J. C. Koehler, J. Weiske, K. Vogeley, N. Koutsouleris, and C. Falter-Wagner. Machine Learning to Study Social Interaction Difficulties in ASD. *Frontiers in Robotics and AI*, 6:132, 2019. doi: 10.3389/frobt.2019.00132

[24] A. L. Georgescu, B. Kuzmanovic, D. Roth, G. Bente, and K. Vogeley. The use of virtual characters to assess and train non-verbal communication in high-functioning autism. *Frontiers in Human Neuroscience*, 8:807, 2014. doi: 10.3389/fnhum.2014.00807

[25] M. González-Franco, D. Pérez-Marcos, B. Spanlang, and M. Slater. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE Virtual Reality Conference (VR)*, pp. 111–114, Mar. 2010. ISSN: 2375-5334. doi: 10.1109/VR.2010.5444805

[26] A. Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. "O'Reilly Media, Inc.", Sept. 2019.

[27] M. Hautzinger, F. Keller, and C. Kühner. *Beck Depressions-Inventar: BDI II.Revision*. Harcourt Test Services, Frankfurt am Main, revision ed., 2006.

[28] H. Hodges, C. Fealko, and N. Soares. Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. *Translational Pediatrics*, 9(Suppl 1):S55–S65, Feb. 2020. doi: 10.21037/tp.2019.09.09

[29] M. Hosozawa, A. Sacker, and N. Cable. Timing of diagnosis, depression and self-harm in adolescents with autism spectrum disorder. *Autism*, 25(1):70–78, Jan. 2021. Publisher: SAGE Publications Ltd. doi: 10.1177/1362361320945540

[30] Y. Huang, S. R. Arnold, K.-R. Foley, and J. N. Trollor. Diagnosis of autism in adulthood: A scoping review. *Autism*, 24(6):1311–1327, Aug. 2020. Publisher: SAGE Publications Ltd. doi: 10.1177/1362361320903128

[31] K. Hume, J. R. Steinbrenner, S. L. Odom, K. L. Morin, S. W. Nowell, B. Tomaszewski, S. Szendrey, N. S. McIntyre, S. Yücesoy-Özkan, and M. N. Savage. Evidence-Based Practices for Children, Youth, and Young Adults with Autism: Third Generation Review. *Journal of Autism and Developmental Disorders*, Jan. 2021. doi: 10.1007/s10803-020-04844-2

[32] M. Jiang and Q. Zhao. Learning visual attention to identify people with autism spectrum disorder. In *Proceedings of the ieee international conference on computer vision*, pp. 3267–3276, 2017.

[33] A. Kirby, L. Edwards, D. Sugden, and S. Rosenblum. The development and standardization of the Adult Developmental Co-ordination Disorders/Dyspraxia Checklist (ADC). *Research in Developmental Disabilities*, 31(1):131–139, 2010. doi: 10.1016/j.ridd.2009.08.010

[34] A. Klin, W. Jones, R. Schultz, F. Volkmar, and D. Cohen. Visual Fixation Patterns During Viewing of Naturalistic Social Situations as Predictors of Social Competence in Individuals With Autism. *Archives of General Psychiatry*, 59(9):809, Sept. 2002. doi: 10.1001/archpsyc.59.9.809

[35] J. C. Koehler, A. L. Georgescu, J. Weiske, M. Spangemacher, L. Burghof, P. Falkai, N. Koutsouleris, W. Tschacher, K. Vogeley, and C. M. Falter-Wagner. Brief Report: Specificity of Interpersonal Synchrony Deficits to Autism Spectrum Disorder and Its Potential for Digitally Assisted Diagnostics. *Journal of Autism and Developmental Disorders*, July 2021. doi: 10.1007/s10803-021-05194-3

[36] A. Koirala, Z. Yu, H. Schiltz, A. Van Hecke, B. Armstrong, and Z. Zheng. A Preliminary Exploration of Virtual Reality-Based Visual and Touch Sensory Processing Assessment for Adolescents With Autism Spectrum Disorder. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:619–628, 2021. Conference Name: IEEE Transactions on Neural Systems and Rehabilitation Engineering. doi: 10.1109/TNSRE.2021.3064148

[37] M.-C. Lai and S. Baron-Cohen. Identifying the lost generation of adults with autism spectrum conditions. *The Lancet Psychiatry*, 2(11):1013–1027, Nov. 2015. Publisher: Elsevier. doi: 10.1016/S2215-0366(15)00277-1

[38] A. Leedham, A. R. Thompson, R. Smith, and M. Freeth. 'I was exhausted trying to figure it out': The experiences of females receiving an autism diagnosis in middle to late adulthood. *Autism*, 24(1):135–146, Jan. 2020. Publisher: SAGE Publications Ltd. doi: 10.1177/1362361319853442

[39] S. Lehrl, J. Merz, G. Burkhard, and S. Fischer. *Mehrfachwahl-Wortschatz-*

Intelligenztest: MWT-B. Perimed-Fachbuch-Verlag-Ges., 1989.
[40] W. Liu, M. Li, and L. Yi. Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. *Autism Research*, 9(8):888–898, 2016. doi: 10.1002/aur.1615
[41] W. Liu, X. Yu, B. Raj, L. Yi, X. Zou, and M. Li. Efficient autism spectrum disorder prediction with eye movement: A machine learning framework. In *2015 International conference on affective computing and intelligent interaction (ACII)*, pp. 649–655. IEEE, 2015.
[42] C. Lord, R. Luyster, K. Gotham, and W. Guthrie. *Autism Diagnostic Observation Schedule. 2nd. (ADOS-2) Manual (Part II): Toddler Module.* Western Psychological Services, Torrence, CA, 2012.
[43] C. Lord, S. Risi, P. S. DiLavore, C. Shulman, A. Thurm, and A. Pickles. Autism from 2 to 9 years of age. *Archives of General Psychiatry*, 63(6):694–701, June 2006. doi: 10.1001/archpsyc.63.6.694
[44] B. B. Maddox, E. S. Brodkin, M. E. Calkins, K. Shea, K. Mullan, J. Hostager, D. S. Mandell, and J. S. Miller. The Accuracy of the ADOS-2 in Identifying Autism among Adults with Complex Psychiatric Conditions. *Journal of Autism and Developmental Disorders*, 47(9):2703–2709, Sept. 2017. doi: 10.1007/s10803-017-3188-z
[45] M. Maskey, J. Rodgers, V. Grahame, M. Glod, E. Honey, J. Kinnear, M. Labus, J. Milne, D. Minos, H. McConachie, and J. R. Parr. A Randomised Controlled Feasibility Trial of Immersive Virtual Reality Treatment with Cognitive Behaviour Therapy for Specific Phobias in Young People with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49(5):1912–1927, May 2019. doi: 10.1007/s10803-018-3861-x
[46] M. Maskey, J. Rodgers, B. Ingham, M. Freeston, G. Evans, M. Labus, and J. R. Parr. Using Virtual Reality Environments to Augment Cognitive Behavioral Therapy for Fears and Phobias in Autistic Adults. *Autism in Adulthood*, 1(2):134–145, June 2019. doi: 10.1089/aut.2018.0019
[47] J. Moon and F. Ke. Exploring the treatment integrity of virtual reality-based social skills training for children with high-functioning autism. *Interactive Learning Environments*, 0(0):1–15, May 2019. doi: 10.1080/10494820.2019.1613665
[48] L. Mottron and D. Bzdok. Autism spectrum heterogeneity: fact or artifact? *Molecular Psychiatry*, 25(12):3178–3185, Dec. 2020. doi: 10.1038/s41380-020-0748-y
[49] D. Neumann, M. L. Spezio, J. Piven, and R. Adolphs. Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, 1(3):194–202, Dec. 2006. doi: 10.1093/scan/nsl030
[50] N. Newbutt, C. Sung, H. J. Kuo, and M. J. Leahy. The potential of virtual reality technologies to support people with an autism condition: A case study of acceptance, presence and negative effects. *Annual Review of Cyber Therapy and Telemedicine (ARCTT)*, 14, Mar. 2016.
[51] S. Optical. Original Stereo Fly Stereotest.
[52] W. H. Organization. ICD-11, 2019.
[53] J. Orlosky, Y. Itoh, M. Ranchet, K. Kiyokawa, J. Morgan, and H. Devos. Emulation of Physician Tasks in Eye-Tracked Virtual Reality for Remote Diagnosis of Neurodegenerative Disease. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1302–1311, Apr. 2017. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2017.2657018
[54] E. A. Papagiannopoulou, K. M. Chitty, D. F. Hermens, I. B. Hickie, and J. Lagopoulos. A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social neuroscience*, 9(6):610–632, 2014. Publisher: Taylor & Francis.
[55] M. Penner, E. Anagnostou, L. Y. Andoni, and W. J. Ungar. Systematic review of clinical guidance documents for autism spectrum disorder diagnostic assessment in select regions. *Autism: The International Journal of Research and Practice*, 22(5):517–527, July 2018. doi: 10.1177/1362361316685879
[56] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *arXiv:1711.05225 [cs, stat]*, Dec. 2017. arXiv: 1711.05225.
[57] C. R. Ramachandiran, N. Jomhari, S. Thiyagaraja, and M. Maria. Virtual Reality Based Behavioural Learning for Autistic Children. *undefined*, 2015.
[58] D. Roth, M. Jording, T. Schmee, P. Kullmann, N. Navab, and K. Vogeley. Towards Computer Aided Diagnosis of Autism Spectrum Disorder Using Virtual Environments. In *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 115–122. IEEE, Utrecht, Netherlands, Dec. 2020. doi: 10.1109/AIVR50618.2020.00029
[59] D. Roth, P. Kullmann, G. Bente, D. Gall, and M. E. Latoschik. Effects of Hybrid and Synthetic Social Gaze in Avatar-Mediated Interactions. In *Adjunct Proceedings of the IEEE International Symposium for Mixed and Augmented Reality 2018*. IEEE, Munich, Germany, 2018. doi: 10.1109/ISMAR-Adjunct.2018.00044
[60] D. Roth and M. E. Latoschik. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3546–3556, Dec. 2020. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2020.3023603
[61] D. Roth, M. E. Latoschik, K. Vogeley, and G. Bente. Hybrid Avatar-Agent Technology – A Conceptual Step Towards Mediated "Social" Virtual Reality and its Respective Challenges. *i-com*, 14(2):107–114, Aug. 2015. Publisher: Oldenbourg Wissenschaftsverlag Section: i-com. doi: 10.1515/icom-2015-0030
[62] D. Roth, J.-L. Lugrin, J. Büser, G. Bente, A. Fuhrmann, and M. E. Latoschik. A simplified inverse kinematic approach for embodied VR applications. In *2016 IEEE Virtual Reality (VR)*, pp. 275–276, Mar. 2016. ISSN: 2375-5334. doi: 10.1109/VR.2016.7504760
[63] D. Roth, J.-P. Stauffert, and M. E. Latoschik. Avatar Embodiment, Behavior Replication, and Kinematics in Virtual Reality. In *VR Developer Gems*. A K Peters/CRC Press, 2019. Num Pages: 26.
[64] M. Slater and A. Steed. A Virtual Presence Counter. *Presence*, 9:413–434, Oct. 2000. doi: 10.1162/105474600566925
[65] H. J. Smith and M. Neff. Communication Behavior in Embodied Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. Association for Computing Machinery, New York, NY, USA, Apr. 2018.
[66] J. Wade, L. Zhang, D. Bian, J. Fan, A. Swanson, A. Weitlauf, M. Sarkar, Z. Warren, and N. Sarkar. A Gaze-Contingent Adaptive Virtual Reality Driving Environment for Intervention in Individuals with Autism Spectrum Disorders. *ACM Transactions on Interactive Intelligent Systems*, 6(1):3:1–3:23, Mar. 2016. doi: 10.1145/2892636
[67] R. Weiß. *Grundintelligenztest Skala 2—Revision (CFT 20-R) [Culture Fair Intelligence Test 20-R—Scale 2].* Jan. 2006.
[68] World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*, 310(20):2191–2194, Nov. 2013. doi: 10.1001/jama.2013.281053
[69] V. Yaneva, S. Eraslan, Y. Yesilada, and R. Mitkov. Detecting high-functioning autism in adults using eye tracking and machine learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(6):1254–1261, 2020. Publisher: IEEE.
[70] V. Yaneva, L. A. Ha, S. Eraslan, Y. Yesilada, and R. Mitkov. Detecting Autism Based on Eye-Tracking Data from Web Searching Tasks. In *Proceedings of the 15th International Web for All Conference*, W4A '18, pp. 1–10. Association for Computing Machinery, New York, NY, USA, Apr. 2018. doi: 10.1145/3192714.3192819
[71] H. Zhao, Z. Zheng, A. Swanson, A. Weitlauf, Z. Warren, and N. Sarkar. Design of a Haptic-Gripper Virtual Reality System (Hg) for Analyzing Fine Motor Behaviors in Children with Autism. *ACM Transactions on Accessible Computing*, 11(4):19:1–19:21, Nov. 2018. doi: 10.1145/3231938
[72] Z. Zheng, Q. Fu, H. Zhao, A. R. Swanson, A. S. Weitlauf, Z. E. Warren, and N. Sarkar. Design of an Autonomous Social Orienting Training System (ASOTS) for Young Children With Autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(6):668–678, June 2017. doi: 10.1109/TNSRE.2016.2598727

Acknowledgements

I want to thank everyone who supported me during this thesis project.

To my supervisors, Prof. Dr. Christine M. Falter-Wagner and Prof. Dr. Lorena Chanes, for their guidance and support, and for providing countless opportunities to grow as a researcher and as a person, becoming more critical, curious, and persistent. Experiencing diverse academic environments and research cultures at LMU and UAB has broadened my perspective and deepened my understanding of research. Thank you for opening the doors of your labs.

To Prof. Dr. Zhuanghua Shi and Prof. Dr. Nikolaos Koutsouleris, my second and third TAC members, for their valuable advice throughout this journey. To Prof. Dr. Daniel Roth, with whom I shared some challenging situations, for his endless motivation and encouragement. To Dr. Enric Duran-Tauleria and Dr. Amaia Hervás, for their incessant collaboration and help.

To the autistic individuals who participated in the studies and to their families: without you, there would have been no research. Learning from you has always been my most valuable lesson.

To my parents, parents-in-law, family, and friends for their constant encouragement and support. To my husband Harry, for his unconditional love and support throughout this journey, and to my best friend Anna, who was always there to listen and help.

Finally, to my colleagues from the NEVIA and the CAS Labs, for many coffee breaks and creating an enjoyable environment at the office.