Título: Percepción y producción de señales no verbales en el autismo: de las predicciones sobre expresiones faciales a la evaluación objetiva de la mirada mediante inteligencia artificial y realidad virtual.

Resumen: La comunicación no verbal desempeña un papel fundamental en las interacciones sociales, al permitir transmitir e interpretar información mediante señales como las expresiones faciales, la mirada o los movimientos corporales. En el autismo se han descrito diferencias tanto en la percepción como en la producción de estas señales. Sin embargo, su caracterización continúa dependiendo en gran medida de medidas clínicas y observacionales, por lo que resulta necesario avanzar hacia procedimientos que permitan cuantificar objetivamente el comportamiento social. Este trabajo investiga la percepción y producción de señales no verbales en personas autistas adultas mediante tres estudios complementarios que combinan el análisis de predicciones sobre expresiones faciales, realidad virtual (RV) y técnicas de inteligencia artificial y aprendizaje automático. Participaron 143 adultos: 74 personas autistas y 69 controles. El primer estudio (34 autistas, 34 controles) examinó las predicciones sobre las expresiones faciales de otras personas y su influencia posterior en la evaluación social. El segundo (6 autistas, 13 controles) consistió en un estudio piloto desarrollado en RV para evaluar la viabilidad del registro objetivo y simultáneo de patrones de mirada y movimientos de cabeza y cuerpo durante una interacción social virtual. A partir de este estudio se desarrolló un paradigma posteriormente validado en una muestra independiente y más amplia en el tercer estudio (40 autistas, 35 controles), incorporando técnicas de aprendizaje automático mediante algoritmos de clasificación supervisada y validación cruzada estratificada y anidada de cinco particiones. Los resultados del primer estudio mostraron que las personas autistas realizaban predicciones sobre expresiones faciales menos estereotípicas que los controles ($B = -0{,}39$, $p = 0{,}002$), aunque ambos grupos utilizaban dichas predicciones de manera comparable al realizar evaluaciones sociales posteriores ($p = 0{,}801$). El estudio piloto confirmó la viabilidad de registrar objetivamente señales no verbales durante interacciones sociales en RV. En el tercer estudio, el algoritmo Gradient Boosting alcanzó una exactitud de 0,80, $F1 = 0{,}84$ y una sensibilidad de 0,95 para la clasificación entre participantes autistas y controles. Los análisis SHAP identificaron los patrones de mirada como los principales predictores de la clasificación, seguidos por los movimientos de cabeza y cuerpo. En conjunto, los tres estudios muestran tanto diferencias como procesos compartidos en la comunicación no verbal en el autismo. Asimismo, la validación del paradigma desarrollado inicialmente en el estudio

piloto respalda la combinación de RV e inteligencia artificial como una aproximación prometedora para obtener medidas cuantitativas y objetivas del comportamiento social, complementarias a la evaluación convencional, y explorar su potencial futuro como biomarcadores digitales del autismo.

Palabras clave: autismo; comunicación no verbal; procesamiento predictivo; realidad virtual; aprendizaje automático.

1. Introducción

El autismo es una condición del neurodesarrollo caracterizada por diferencias persistentes en la comunicación e interacción social y por patrones restringidos o repetitivos de comportamiento, intereses o actividades (American Psychiatric Association [APA], 2022). La forma y la intensidad con la que se manifiestan sus características varían considerablemente entre personas autistas[1] (Kim et al., 2024a) y pueden verse moduladas por factores como el sexo y el género, las capacidades cognitivas, las habilidades sociales, el momento de identificación y las estrategias aprendidas a lo largo del desarrollo (APA, 2022). Dentro de esta heterogeneidad, la comunicación no verbal constituye un ámbito de especial relevancia. Las señales no verbales son fundamentales para el desarrollo y mantenimiento de las relaciones sociales y contribuyen a transmitir estados afectivos, actitudes e intenciones. Incluyen, entre otras modalidades, las expresiones faciales, la mirada interpersonal, los movimientos de la cabeza, el cuerpo y las manos, los gestos, la postura y orientación corporal, la distancia interpersonal y la sincronía entre interlocutores. Las diferencias en la comunicación no verbal forman parte de la propia caracterización clínica del autismo (APA, 2022). El DSM-5-TR incluye entre ellas un uso ausente, reducido o atípico del contacto visual, los gestos, las expresiones faciales, la orientación corporal o la entonación, así como diferencias en la coordinación entre componentes verbales y no verbales de la comunicación (APA, 2022). Sin embargo, la considerable variabilidad existente dentro del espectro hace que una caracterización global en términos de “déficit” resulte insuficiente para explicar cómo las personas autistas

[1] Nota sobre terminología. A lo largo de este trabajo se emplea preferentemente la expresión «persona autista», siguiendo un enfoque de lenguaje centrado en la identidad (*identity-first language*). Esta elección se basa en estudios que han mostrado una preferencia general por esta terminología entre personas autistas adultas (Botha et al., 2022; Bury et al., 2020, 2023; Geelhand et al., 2023; Kapp et al., 2013; Keating et al., 2023; Kenny et al., 2015; Lei et al., 2021; Taboas et al., 2022). No obstante, no existe un consenso universal y las preferencias terminológicas pueden variar entre personas, colectivos, contextos culturales e idiomas. Por ello, esta elección no pretende establecer una terminología universalmente preferible, sino adoptar la opción que refleja la preferencia mayoritaria encontrada en los estudios considerados, reconociendo y respetando la diversidad de preferencias individuales.

perciben, interpretan y producen las diferentes señales no verbales. Por este motivo, el presente trabajo investiga estas señales tanto a nivel perceptivo como productivo.

Percepción de las expresiones faciales en el autismo

Una parte importante de la investigación sobre percepción social en el autismo se ha centrado en el reconocimiento de expresiones faciales emocionales, tradicionalmente asumiendo que determinadas configuraciones faciales corresponden a categorías emocionales específicas y relativamente universales. Numerosos paradigmas han evaluado así la capacidad para asignar expresiones prototípicas a categorías como alegría, tristeza, miedo o enfado. La evidencia muestra diferencias entre personas autistas y controles en estas tareas, aunque el rendimiento varía según la complejidad e intensidad de los estímulos y el tipo de procesamiento requerido (Barrett et al., 2019; Keating & Cook, 2021; Yeung, 2022). Sin embargo, evaluar las respuestas como correctas o incorrectas presupone una correspondencia normativa entre situación, emoción y configuración facial. Una alternativa consiste en estudiar las predicciones de la propia persona perceptora, examinando hasta qué punto la expresión observada coincide con la esperada en un determinado contexto. Esta perspectiva se relaciona con los modelos de procesamiento predictivo, según los cuales el cerebro genera continuamente predicciones basadas en la experiencia previa y las contrasta con la información sensorial entrante (Friston, 2005, 2010; Yon & Frith, 2021). Las discrepancias generan errores de predicción que contribuyen a actualizar los modelos internos que guían la percepción y la conducta. En el ámbito emocional, estas ideas han sido incorporadas a perspectivas construccionistas que atribuyen un papel fundamental a la experiencia previa y las expectativas en la percepción de las expresiones emocionales (Barrett, 2017; Hoemann et al., 2020). En el autismo, diversos modelos de procesamiento predictivo proponen diferencias en la generación, ponderación o actualización de las predicciones y los errores de predicción (Cannon et al., 2021; Lawson et al., 2014; Pellicano & Burr, 2012; Van de Cruys et al., 2014). En el ámbito social, este marco permite estudiar no solo la identificación normativa de expresiones faciales, sino también qué expresión se espera en un contexto y hasta qué punto la observada coincide con esa expectativa. Estas predicciones pueden influir además en la evaluación social: en población no clínica, las expresiones que coinciden mejor con las expectativas del observador se consideran más agradables y confiables, lo que sugiere un papel activo de las predicciones en la percepción y evaluación social (Chanes et al., 2018).

Esta aproximación plantea dos preguntas: si las personas autistas generan predicciones sobre las expresiones faciales con el mismo grado de estereotipicidad que los controles y si utilizan sus propias predicciones de manera diferente al evaluar socialmente a otras personas. Esta distinción permite separar la naturaleza de la predicción de su uso: una persona puede mantener expectativas distintas de las normativamente estereotípicas y, al mismo tiempo, utilizarlas funcionalmente para interpretar y evaluar a los demás. Esta fue la cuestión central del primer estudio, que se alejó de paradigmas basados exclusivamente en “aciertos” y “errores” para examinar la correspondencia entre las expectativas individuales y la información social observada.

Producción de señales no verbales: mirada y movimiento corporal

La comunicación no verbal no depende únicamente de cómo se perciben las señales de otras personas. Durante una interacción, cada interlocutor produce continuamente señales mediante la mirada, la orientación de la cabeza, la postura, los gestos y los movimientos corporales, que contribuyen al establecimiento de la atención conjunta, la regulación de la interacción y la interpretación de las intenciones sociales (Hamilton, 2016; Hietanen, 2018; Stephenson et al., 2021). Estas conductas desempeñan, además, un papel relevante desde etapas tempranas del desarrollo social y comunicativo (Farroni et al., 2002; Mundy et al., 2003).

La mirada ha recibido especial atención en la investigación sobre autismo. Estudios mediante eye-tracking (del inglés, seguimiento de la mirada) han descrito diferencias en la atención visual hacia estímulos sociales, especialmente hacia la región ocular (Camero et al., 2021; Klin et al., 2002; Papagiannopoulou et al., 2014; Riddiford et al., 2022). Se ha observado una menor fijación en los ojos y mayor orientación hacia otras regiones faciales, como la boca, el cuerpo o elementos no sociales del entorno (Camero et al., 2021; Kim et al., 2024b; Klin et al., 2002). Además, una menor orientación hacia la cara, la cabeza y los ojos se ha asociado con mayores dificultades sociales y características autistas más pronunciadas (Riddiford et al., 2022). También se han descrito diferencias en los movimientos de cabeza y cuerpo, incluyendo mayor movimiento de cabeza, variaciones en su velocidad y dirección y menor complejidad del movimiento corporal durante interacciones sociales (Campbell et al., 2018; Dawson et al., 2018; Martin et al., 2018; Zhao et al., 2021, 2022), así como diferencias en movimientos de las manos y parámetros cinemáticos durante tareas de imitación y motoras (Ardalan et al., 2019; Li et al., 2017).

Estos hallazgos han llevado a estudiar la mirada y otras señales no verbales como posibles marcadores biológicos conductuales o digitales del autismo (Parellada et al., 2022; Wei et al., 2023). Sin embargo, los resultados varían según el área facial, la persona y el contexto. Mientras que la menor fijación ocular aparece de forma relativamente consistente, las diferencias en la atención hacia la boca dependen más del contexto social y emocional y no siempre alcanzan significación estadística (Papagiannopoulou et al., 2014; Riddiford et al., 2022; Wieckowski & White, 2017). Esta variabilidad subraya la heterogeneidad del autismo y la influencia del contexto experimental. Además, los paradigmas tradicionales presentan un compromiso entre control y validez ecológica: los estímulos estáticos o pregrabados facilitan la estandarización, pero reproducen de forma limitada la naturaleza dinámica y recíproca de la interacción social, mientras que las interacciones naturalistas ofrecen mayor realismo a costa de introducir variabilidad. Esta tensión constituye un reto metodológico central en la investigación social y clínica (Parsons, 2015).

Realidad virtual para el estudio objetivo del comportamiento social

La realidad virtual (RV) permite abordar esta tensión combinando un elevado control experimental con entornos interactivos y socialmente significativos (Parsons, 2015). Permite presentar escenarios e interacciones estandarizados y registrar de forma continua y objetiva el comportamiento en condiciones reproducibles (Alcañiz et al., 2022; Savickaite et al., 2022). Su uso en investigación sobre autismo ha aumentado, especialmente en intervenciones sobre habilidades sociales, regulación emocional, fobias y entrenamiento motor (Li et al., 2023; Lorenzo et al., 2016; Maskey et al., 2014, 2019; Zhao et al., 2018). Sin embargo, su aplicación a la evaluación y caracterización diagnóstica sigue menos desarrollada, pese a su potencial para registrar de forma controlada y objetiva el comportamiento social (Savickaite et al., 2022).

En este trabajo se desarrolló una situación inmersiva que simulaba una compra en un supermercado, en la que la persona interactuaba con un empleado virtual mientras se registraban objetivamente señales no verbales como la mirada y los movimientos de cabeza y cuerpo. El estudio piloto confirmó la viabilidad del sistema y detectó diferencias entre personas autistas y controles, especialmente en la fijación sobre la región ocular, la orientación hacia el fondo y determinados parámetros de movimiento de la cabeza. Esta aproximación se alinea con estudios que combinan RV, seguimiento ocular y aprendizaje automático para identificar patrones conductuales relacionados con el autismo (Alcañiz et

al., 2020, 2022; Roth et al., 2020). En conjunto, estos trabajos respaldan el potencial de la RV para estudiar señales no verbales de forma objetiva y reproducible, manteniendo mayor realismo social que los paradigmas basados en estímulos estáticos.

Inteligencia artificial y marcadores digitales

El registro continuo de la mirada y el movimiento genera datos multidimensionales difíciles de resumir mediante una única medida. En este contexto, la inteligencia artificial (IA), especialmente el aprendizaje automático (machine learning), permite identificar patrones distribuidos entre distintas señales conductuales. Esta aproximación resulta especialmente relevante en la edad adulta. Aunque el autismo puede identificarse de forma fiable desde edades tempranas (Lord et al., 2006), algunas personas permanecen sin identificar hasta la edad adulta (Lai & Baron-Cohen, 2015). La heterogeneidad de su presentación, las diferencias asociadas al sexo y al género y las estrategias aprendidas para desenvolverse socialmente pueden dificultar o retrasar su reconocimiento y contribuir a su variabilidad (Hull et al., 2020; Lai & Szatmari, 2020). La evaluación puede ser particularmente compleja ante condiciones psiquiátricas con características solapadas. Aunque herramientas estandarizadas como el Autism Diagnostic Observation Schedule–Second Edition (ADOS-2; Lord et al., 2012) son relevantes, su capacidad discriminativa puede reducirse en adultos con presentaciones psiquiátricas complejas. Maddox et al. (2017), por ejemplo, observaron una elevada tasa de falsos positivos, subrayando que sus puntuaciones no deben interpretarse aisladamente del conjunto de la evaluación clínica.

El estudio piloto que se presenta en este trabajo mostró que un sistema de RV podía registrar señales no verbales de forma controlada y utilizarlas para la clasificación de patrones. Sin embargo, las elevadas tasas de clasificación obtenidas con una muestra pequeña requerían validación en una muestra mayor. El tercer estudio responde a esta necesidad utilizando la muestra completa de 75 participantes y un procedimiento metodológico más robusto para evaluar la capacidad de un algoritmo de aprendizaje automático para diferenciar patrones multimodales de comportamiento no verbal entre personas autistas y controles.

Objetivos de la investigación

El objetivo general de esta investigación fue caracterizar la percepción y producción de señales de comunicación no verbal en personas autistas adultas mediante paradigmas conductuales, realidad virtual y técnicas de inteligencia artificial.

El objetivo del Estudio 1, titulado en este trabajo como “Predicciones sobre expresiones faciales y evaluación social”, fue examinar las predicciones sobre las expresiones faciales y su influencia en la evaluación social. Se planteó que las personas autistas presentarían predicciones menos estereotípicas que los controles (H1) y que estas diferencias podrían asociarse con una utilización diferente de las predicciones en la evaluación social (H2). El Estudio 2, titulado “Estudio piloto mediante realidad virtual”, tuvo como objeto evaluar la viabilidad de registrar objetivamente señales de comunicación no verbal durante una interacción social inmersiva y explorar su utilización mediante aprendizaje automático. Se planteó que las personas autistas y los controles mostrarían diferencias en los patrones de fijación sobre los ojos, la boca y el fondo (H3) y que las señales de mirada y los movimientos de cabeza y manos permitirían explorar la clasificación entre ambos grupos mediante aprendizaje automático (H4). Finalmente, en el Estudio 3, “Evaluación en la muestra completa mediante inteligencia artificial” se amplió el estudio piloto y se evaluó en la muestra completa la capacidad de las señales no verbales para diferenciar entre personas autistas y controles. Se planteó que la combinación de características de mirada y movimientos de cabeza y manos permitiría clasificar entre ambos grupos mediante aprendizaje automático (H5). Además, se exploró qué características contribuían en mayor medida a la clasificación mediante valores SHAP.

2. Método

Diseño general

En total, este trabajo incluyó a 143 participantes adultos: 74 personas autistas y 69 controles. La investigación integra tres estudios complementarios. El Estudio 1 contó con una muestra independiente de 68 participantes. Los Estudios 2 y 3 corresponden a una misma línea de investigación basada en RV: el Estudio 2 analizó inicialmente una submuestra de 19 participantes como estudio piloto y el Estudio 3 extendió posteriormente el procedimiento a la muestra completa de 75 participantes.

Estudio 1. Predicciones sobre expresiones faciales y evaluación social

Participantes

Participaron 68 personas adultas, distribuidas en dos grupos: 34 personas autistas y 34 controles. El grupo autista estuvo formado por 16 mujeres y 18 hombres, con una edad media de 28 años (DE = 10), mientras que el grupo control estuvo compuesto igualmente por 16 mujeres y 18 hombres, con una edad media de 31 años (DE = 7). La diferencia de edad entre grupos fue pequeña, aunque estadísticamente significativa (U = 754,5, p = 0,031). El tamaño de la muestra se estableció tomando como referencia investigaciones previas que habían empleado el mismo paradigma en población adulta sin condiciones clínicas (Chanes et al., 2018), así como estudios previos sobre percepción de expresiones faciales en autismo (véase, por ejemplo, Yeung, 2022).

Las personas del grupo autista fueron reclutadas en el Instituto Global de Atención Integral del Neurodesarrollo (IGAIN), en Barcelona, y contaban con un diagnóstico clínico de autismo previamente establecido. La información relativa al diagnóstico, las condiciones coexistentes y el tratamiento farmacológico en el momento de la participación se obtuvo a partir de la historia clínica. La presencia de otras condiciones de salud mental o el uso de medicación no constituyeron criterios de exclusión en el grupo autista, con el objetivo de mantener una muestra clínicamente representativa. En concreto, 21 participantes autistas (62%) presentaban al menos una condición coexistente: 13 (38%) participantes presentaban una y 8 (24%) dos o más. Entre las condiciones coexistentes más frecuentes se encontraban los trastornos de ansiedad. Todos los participantes incluidos eran adultos sin discapacidad intelectual. El grupo control fue reclutado mediante anuncios difundidos a través de canales universitarios y redes sociales. Los participantes fueron seleccionados con el objetivo de emparejarlos con las personas del grupo autista en género y edad, estableciendo para esta última un margen de ± 5 años. Se aplicaron los siguientes criterios de inclusión: (i) edad ≥18 años; (ii) ausencia de antecedentes de lesión cerebral u otras condiciones neurológicas; (iii) ausencia de un diagnóstico actual o previo de ninguna condición psiquiátrica, ni encontrarse en tratamiento con medicación psicoactiva; (iv) visión normal o corregida a normal; y (v) ser hablante nativo de español o bilingüe catalán-español.

Como medida adicional de caracterización, se administró la versión abreviada del Autism-Spectrum Quotient (AQ-Short; Hoekstra et al., 2011), utilizando su adaptación española (Lugo-Marín et al., 2019). El AQ-Short consta de 28 ítems autoinformados y

permite cuantificar características asociadas al autismo. Como era esperable, las personas autistas presentaron puntuaciones significativamente superiores (M = 72, DE = 11) a las del grupo control (M = 53, DE = 10; U = 110, p < 0,001).

Procedimiento y paradigma experimental

Tras la recogida de información sociodemográfica, los participantes realizaron una tarea conductual de procesamiento predictivo y percepción social desarrollada previamente por Chanes et al. (2018). Posteriormente completaron el AQ-Short (Hoekstra et al., 2011; versión española: Lugo-Marín et al., 2019), junto con otros cuestionarios que no fueron objeto de los análisis del presente estudio. La tarea había sido utilizada previamente para investigar cómo las expectativas sobre las expresiones faciales influyen en la evaluación social (Chanes et al., 2018; Draganov et al., 2023) y posteriormente se había aplicado al estudio del procesamiento predictivo en depresión (Ramos-Grille et al., 2022). En investigaciones previas con esta tarea se había observado que, en población adulta control, las expresiones faciales estereotípicamente congruentes con el contexto emocional son consideradas más similares a las predichas que las expresiones incongruentes y, crucialmente, que las expresiones percibidas como más predecibles se asocian con evaluaciones sociales más favorables (Chanes et al., 2018). Los estímulos se presentaron mediante E-Prime 3 en un ordenador portátil con pantalla de 14 pulgadas y resolución de 1920 × 1080 píxeles. En cada ensayo se presentaba inicialmente la fotografía de una persona junto con un breve escenario escrito que describía una situación normativamente asociada con una de tres categorías emocionales: alegría, tristeza o miedo. Por ejemplo, el escenario podía describir una situación habitualmente relacionada con tristeza, como asistir a un funeral, o con alegría, como disfrutar de un momento con amigos. Se pedía al participante que, mientras leía el escenario, imaginara cómo se vería la persona mostrada en esa situación. A continuación se presentaba nuevamente el rostro de la misma persona mostrando una expresión facial estereotípica correspondiente a alegría, tristeza o miedo. La expresión presentada podía coincidir con la emoción normativamente evocada por el escenario —condición congruente (matched), por ejemplo, una expresión triste después de un escenario de tristeza— o corresponder a una emoción diferente —condición incongruente (non-matched), por ejemplo, una expresión triste después de un escenario de alegría o miedo—. Tras observar la expresión facial, los participantes realizaban dos valoraciones consecutivas mediante escalas de cuatro puntos y sin límite de tiempo. En primer lugar, indicaban hasta qué punto el rostro observado se parecía a lo que habían imaginado al leer el escenario, desde 1 («nada

similar») hasta 4 («muy similar»). Esta medida constituyó la puntuación de predictibilidad y se utilizó como aproximación conductual al grado de correspondencia entre la predicción individual y la información finalmente observada. En segundo lugar, evaluaban hasta qué punto les resultaba agradable la persona, desde 1 («muy desagradable») hasta 4 («muy agradable»), proporcionando así la medida de evaluación social o agradabilidad.



Figura 1. Representación esquemática del paradigma experimental utilizado para evaluar las predicciones sobre expresiones faciales y su influencia en la evaluación social (extraído de Robles et al., 2024).

Un aspecto metodológico especialmente importante es que en ningún momento se mencionaron explícitamente a los participantes las categorías emocionales. La tarea no les pedía identificar si una expresión representaba alegría, tristeza o miedo ni seleccionar una respuesta considerada correcta. Su objetivo era determinar en qué medida la expresión presentada coincidía con su propia predicción sobre cómo se vería esa persona en la situación descrita. Por este motivo, la tarea no generaba respuestas correctas o incorrectas ni se calculó una medida de exactitud. Esta característica permitió estudiar las predicciones individuales sin asumir de antemano que la expresión estereotípicamente asociada a un contexto constituía necesariamente la respuesta que el participante debía esperar.

Antes de comenzar la fase experimental se aplicó un procedimiento específico para asegurar la comprensión de la tarea. Los participantes leyeron las instrucciones en pantalla, que incluían fotografías y ejemplos de las escalas de respuesta, pudieron resolver sus dudas con la persona investigadora y realizaron unos ensayos de prueba.

La tarea experimental constó de 45 ensayos, distribuidos en cinco bloques de nueve ensayos, con posibilidad de realizar una pausa breve entre bloques. Se incluyeron 15 escenarios para cada una de las tres categorías emocionales —alegría, tristeza y miedo—. De los 15 ensayos correspondientes a cada emoción, nueve presentaban posteriormente la expresión facial estereotípicamente congruente y seis una expresión

correspondiente a una de las otras dos emociones. En total, cada participante realizó 27 ensayos congruentes y 18 incongruentes.

Análisis estadístico

Los datos correspondientes a los ensayos individuales se analizaron mediante modelos lineales jerárquicos (*Hierarchical Linear Modeling*, HLM; HLM 7.0), en lugar de agregar previamente las respuestas de cada participante. Esta aproximación permitió modelizar la variabilidad ensayo a ensayo teniendo en cuenta la estructura anidada de los datos, es decir, múltiples observaciones dentro de cada participante. Los participantes se trataron como factor aleatorio y los parámetros se estimaron mediante máxima verosimilitud restringida (Raudenbush & Bryk, 2002). Todos los modelos incluyeron interceptos aleatorios. Las variables continuas a nivel de ensayo se centraron respecto a la media de cada participante, mientras que las variables categóricas codificadas mediante variables dummy se introdujeron sin centrar (Enders & Tofighi, 2007).

En primer lugar, se examinó si existían diferencias globales entre grupos en las puntuaciones de predictibilidad. Posteriormente se evaluó el efecto de la congruencia de la expresión mediante un modelo en el que la puntuación de predictibilidad de cada ensayo constituyó la variable dependiente, la condición —congruente frente a incongruente— se introdujo como predictor a nivel de ensayo y el grupo —autista frente a control— como predictor a nivel de participante. Se incluyó la interacción grupo × congruencia, que permitió determinar si la diferencia entre expresiones congruentes e incongruentes —el denominado match effect— difería entre ambos grupos. Los análisis posteriores examinaron por separado las condiciones congruentes e incongruentes y las tres categorías emocionales. Finalmente, para determinar si las predicciones influían en la evaluación social, se estimó un modelo HLM en el que la puntuación de agradabilidad de cada ensayo fue la variable dependiente, la predictibilidad se introdujo como predictor a nivel de ensayo y el grupo como predictor a nivel de participante. La interacción predictibilidad × grupo permitió evaluar si la relación entre la correspondencia con las expectativas individuales y la agradabilidad atribuida a la persona difería entre participantes autistas y controles.

Como análisis complementario, se exploró si el grado de estereotipicidad de las predicciones se relacionaba dimensionalmente con las características autistas en personas sin diagnóstico de autismo. Para este análisis específico se incorporaron siete participantes control adicionales procedentes de otro estudio con los mismos criterios de

inclusión, obteniéndose una muestra de 41 personas. Se calculó para cada participante el match effect y se examinó su asociación con las puntuaciones del AQ-Short mediante regresión lineal.

Estudio 2. Estudio piloto mediante realidad virtual

Participantes

El estudio piloto incluyó 19 participantes adultos, distribuidos en dos grupos: 6 personas autistas y 13 controles. Se empleó un diseño entre grupos para comparar las señales de comunicación no verbal registradas durante una interacción social en realidad virtual. El grupo de personas autistas estuvo formado por participantes con diagnóstico de autismo. El grupo control estuvo constituido por personas sin diagnósticos psiquiátricos actuales o previos y sin tratamiento con medicación psicoactiva. También se aplicaron los criterios de inclusión y exclusión del Estudio 1, si bien eran hablantes nativos del alemán, ya que se reclutaron en el Hospital de la Universidad LMU de Múnich.

Procedimiento y paradigma experimental

Antes de la tarea experimental se llevó a cabo una caracterización de los participantes que incluyó la recogida de datos demográficos, pruebas de capacidad intelectual y visual y cuestionarios psicológicos y relacionados con la experiencia de realidad virtual.

Se desarrolló un sistema inmersivo de realidad virtual de un solo usuario destinado a generar una situación de interacción social controlada y permitir el registro objetivo de señales de comunicación no verbal. El entorno virtual reproducía una actividad cotidiana: realizar una compra en un supermercado e interactuar con un agente virtual que desempeñaba el papel de vendedor. El sistema utilizó un visor HTC VIVE Pro Eye con seguimiento ocular integrado y dos mandos manuales, junto con tres sensores para el seguimiento espacial. El seguimiento ocular se realizó mediante Tobii XR SDK. Esta configuración permitió registrar simultáneamente la dirección de la mirada y la posición y rotación de la cabeza y de ambas manos durante la experiencia virtual. En total, el sistema registraba 27 características relacionadas con la mirada y el movimiento corporal.

Antes de comenzar la tarea experimental, los participantes realizaron una fase específica de familiarización con la tecnología de realidad virtual. Para ello se desarrolló un espacio virtual introductorio con una apariencia similar a la entrada de un mercado.

Esta fase permitía familiarizarse con el entorno inmersivo y aprender a utilizar correctamente los mandos y los mecanismos necesarios para interactuar posteriormente con los objetos del supermercado. La tarea experimental consistía en realizar una compra dentro de un supermercado virtual. Los participantes debían adquirir cinco productos, presentados en el mismo orden para todos ellos, mientras interactuaban con un agente virtual que representaba a un vendedor del establecimiento. Los participantes utilizaban los mandos del sistema para seleccionar los productos. Cuando seleccionaban un artículo incorrecto, el vendedor les indicaba que debían intentarlo de nuevo. Cuando seleccionaban correctamente uno de los productos de la lista, el agente virtual acercaba el artículo al participante y comenzaba una interacción social relacionada con dicho producto. Después de cada selección correcta, el vendedor proporcionaba información neutra relacionada con el producto durante aproximadamente 90 segundos de media. Estas intervenciones constituían los periodos de interacción social de especial interés para el estudio. Mientras el vendedor hablaba, se registraron las señales no verbales realizadas espontáneamente por los participantes. Tras completar la compra de los cinco ítems, el agente virtual se despedía y finalizaba la tarea.

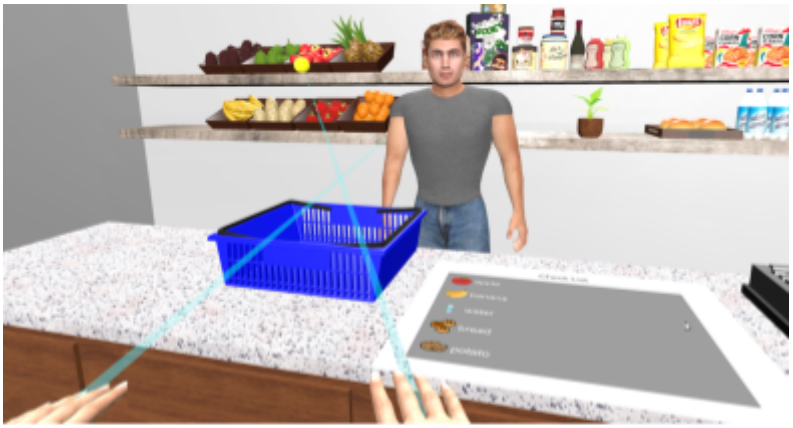


Figura 2. Entorno de realidad virtual que simula un supermercado. El agente (detrás del mostrador) espera a que el participante elija uno de los artículos de la lista que se encuentran sobre el mostrador (extraído de Robles et al., 2022).

El agente virtual fue diseñado para generar una interacción social visual y verbal con el participante. Durante sus intervenciones podía dirigir la mirada hacia este y realizar movimientos coordinados de ojos, cabeza y parte superior del cuerpo. Se incorporaron animaciones de ojos y párpados, parpadeo, movimientos de cabeza y sincronización del movimiento de la boca con el habla. La interacción verbal se presentó mediante grabaciones de voz humana natural adaptadas al contenido de las intervenciones del personaje.

Registro de la mirada y del movimiento corporal

El sistema registró de forma continua la dirección de la mirada durante los periodos de interacción social. Para analizar hacia dónde dirigían la atención los participantes se definieron diferentes áreas de interés (Areas of Interest, AOIs) asociadas al agente virtual. Las áreas de interés incluían regiones socialmente relevantes correspondientes a los ojos, la boca y la cabeza. Estas AOIs eran dinámicas y estaban vinculadas a puntos de referencia del agente virtual, de forma que se desplazaban junto con él durante la interacción. Cuando el vector de mirada del participante intersectaba una de estas regiones, el sistema identificaba automáticamente el área observada. El resto del entorno virtual se consideró fondo (background), permitiendo comparar la atención dirigida hacia regiones sociales del agente con la dirigida hacia otros elementos del escenario. A partir de estos registros se obtuvieron medidas relacionadas con la fijación de la mirada sobre las diferentes áreas de interés, que posteriormente permitieron comparar los patrones de atención visual entre personas autistas y controles. Asimismo, se registraron continuamente la posición y la rotación de la cabeza y de ambas manos. Estas señales permitieron obtener características relacionadas con el desplazamiento y los cambios de orientación de cada uno de estos elementos en el espacio tridimensional.



Figura 3. Regiones dinámicas de interés en 3D (ojos, rostro, boca) del agente, a partir de las cuales se registran y clasifican los datos de seguimiento ocular relativos al foco de la mirada del participante (extraído de Robles et al., 2022).

Análisis estadístico y clasificación mediante aprendizaje automático

Las diferencias entre las personas autistas y los controles se evaluaron mediante comparaciones estadísticas de las características extraídas. Estos análisis permitieron examinar de forma independiente si los patrones de fijación visual, desplazamiento de la mirada y movimiento de la cabeza diferían entre ambos grupos durante los periodos estandarizados de interacción con el agente virtual. Los movimientos de las manos también fueron registrados y se incorporaron a los análisis exploratorios de clasificación de patrones.

Además de las comparaciones estadísticas entre grupos, se examinó si las señales no verbales registradas durante la interacción podían utilizarse para la clasificación automática de personas autistas y controles mediante técnicas de aprendizaje automático. Para ello se emplearon características derivadas de la mirada y de los movimientos de cabeza y manos, explorando diferentes combinaciones de variables como entrada de los modelos de clasificación. El procesamiento de los datos y los análisis computacionales se realizaron en Python mediante herramientas específicas de aprendizaje automático, entre ellas Scikit-learn y TensorFlow. El procedimiento exploró distintas combinaciones de características con el objetivo de determinar si los patrones multimodales de comunicación no verbal registrados durante la tarea contenían información suficiente para discriminar entre ambos grupos. Dado el reducido tamaño de la muestra, los análisis de aprendizaje automático se plantearon con un carácter exploratorio y de viabilidad, y no como una estimación definitiva de la capacidad de generalización de los modelos. Por tanto, el objetivo de esta fase fue determinar si las señales de comunicación no verbal obtenidas mediante el sistema de realidad virtual podían constituir una fuente de información susceptible de ser utilizada para la clasificación de patrones. La evaluación posterior de esta aproximación en una muestra de mayor tamaño y mediante un procedimiento de validación más robusto se llevó a cabo en el Estudio 3.

Estudio 3. Evaluación en la muestra completa mediante inteligencia artificial

Participantes

Participaron 75 adultos, distribuidos en dos grupos: 40 personas autistas y 35 controles. El grupo autista estuvo formado por 19 mujeres y 21 hombres, con una edad media de 30 años (DE = 7), mientras que el grupo control estuvo compuesto por 18 mujeres y 17 hombres, con una edad media de 27 años (DE = 7). Esta muestra incluía a los participantes del Estudio 2 y representaba la ampliación del estudio piloto a la muestra completa. Se mantuvieron los mismos criterios generales de inclusión y exclusión utilizados en el Estudio 2. Antes de la tarea experimental se recogieron datos demográficos y se administraron pruebas de capacidad intelectual y visual, así como cuestionarios psicológicos y relacionados con la experiencia de realidad virtual.

Procedimiento y paradigma experimental

El procedimiento y paradigma experimental es el mismo descrito en el Estudio 2. Los participantes realizaron la misma tarea de realidad virtual de supermercado utilizada

en el Estudio 2, diseñada para simular una situación cotidiana de interacción social. Durante la experiencia, interactuaban con un agente virtual que desempeñaba el papel de vendedor mientras se registraban de forma continua sus patrones de mirada y los movimientos de cabeza y manos. Asimismo, se mantuvieron las AOIs y los datos sobre las conductas no verbales de los participantes fueron registrados de la misma manera que en el Estudio 2.

Análisis mediante aprendizaje automático

El objetivo principal del análisis fue evaluar en la muestra completa si los patrones multimodales de comunicación no verbal identificados mediante la tarea de realidad virtual permitían diferenciar entre personas autistas y controles mediante aprendizaje automático, ampliando y sometiendo a una evaluación metodológicamente más robusta los resultados exploratorios del Estudio 2. Para ello, se aplicaron técnicas de inteligencia artificial y aprendizaje automático sobre las características derivadas de la mirada y de los movimientos de cabeza y manos. El algoritmo evaluado fue un clasificador supervisado basado en *Gradient Boosting*. El entrenamiento y la evaluación del modelo se realizaron mediante un esquema de validación cruzada estratificada y anidada de cinco particiones (en inglés, *stratified nested 5-fold cross-validation*). La estratificación permitió mantener la proporción de participantes autistas y controles en los distintos subconjuntos, mientras que el carácter anidado del procedimiento se utilizó para reducir el riesgo de sobreajuste y obtener una estimación más robusta de la capacidad de generalización del modelo. El rendimiento del modelo se evaluó mediante varias métricas complementarias: exactitud (*accuracy*), sensibilidad, especificidad, precisión y puntuación F1. Esta estrategia permitió caracterizar no solo la proporción global de clasificaciones correctas, sino también la capacidad del algoritmo para identificar correctamente a las personas autistas y a los controles. Además, se examinó la contribución de las distintas características conductuales a las predicciones del modelo mediante valores SHAP (*SHapley Additive exPlanations*). Este procedimiento permitió cuantificar la contribución de cada característica a las predicciones del algoritmo y examinar qué señales no verbales contribuían en mayor medida a la diferenciación entre ambos grupos.

Consideraciones éticas

Los estudios incluidos siguieron los principios de la Declaración de Helsinki y sus protocolos fueron evaluados y aprobados por los correspondientes comités de ética de investigación de la Universidad Autónoma de Barcelona y de la Universidad LMU de

Múnich. Todos los participantes recibieron información sobre los objetivos y procedimientos del estudio, su derecho a rechazar la participación o retirar el consentimiento sin consecuencias y proporcionaron consentimiento informado antes de participar.

3. Resultados

Estudio 1. Predicciones sobre expresiones faciales y evaluación social

Predictibilidad global de las expresiones faciales

En primer lugar, se examinó si existían diferencias globales entre grupos en el grado en que las expresiones faciales observadas coincidían con las predicciones de los participantes. Las personas autistas otorgaron, en promedio, puntuaciones de predictibilidad significativamente inferiores (M = 2,49 ± 0,07) a las del grupo control (M = 2,65 ± 0,04) (B = −0,17 ± 0,08, t(66) = 2,16, p = 0,034). Por tanto, considerando conjuntamente todas las condiciones experimentales, las expresiones faciales observadas coincidieron en menor medida con las expectativas previas de las personas autistas.

Predicciones en función de la congruencia entre el escenario y la expresión facial

Como era esperable, ambos grupos otorgaron puntuaciones de predictibilidad superiores a las expresiones congruentes con la emoción normativamente evocada por el escenario que a las expresiones incongruentes. En el grupo control, las expresiones congruentes fueron valoradas como considerablemente más predecibles (M = 3,25 ± 0,05) que las incongruentes (M = 1,76 ± 0,05) (B = 1,49 ± 0,06, t(66) = 24,85, p < 0,001). El mismo patrón se observó en las personas autistas, que también valoraron las expresiones congruentes (M = 2,93 ± 0,09) como más predecibles que las incongruentes (M = 1,82 ± 0,08) (B = 1,11 ± 0,10, t(66) = 11,05, p < 0,001). Sin embargo, la interacción entre grupo y congruencia fue significativa (B = −0,39 ± 0,12, t(66) = 3,30, p = 0,002). Es decir, la diferencia entre las puntuaciones de predictibilidad de las expresiones congruentes e incongruentes —el denominado match effect— fue significativamente menor en las personas autistas que en los controles.

Los análisis realizados por separado para ambas condiciones permitieron precisar el origen de esta interacción. En los ensayos congruentes, las personas autistas consideraron las expresiones faciales significativamente menos predecibles que los controles (B = 0,32 ± 0,10, t(66) = 3,25, p = 0,002). En cambio, no se observaron

diferencias entre grupos cuando la expresión facial era incongruente con la emoción normativamente asociada al escenario (B = −0,06 ± 0,09, t(66) = 0,69, p = 0,490). En conjunto, este patrón indica que las diferencias entre grupos no se debieron a que las personas autistas valoraran sistemáticamente todas las expresiones como menos predecibles. La diferencia apareció específicamente ante las expresiones estereotípicamente esperables para el contexto: estas coincidieron en menor medida con las predicciones individuales de las personas autistas que con las de los controles.

Predicciones en función de la categoría emocional

A continuación, se examinó si el menor match effect observado en las personas autistas dependía de la emoción evocada por el escenario. El patrón general se reprodujo en las tres categorías emocionales estudiadas —alegría, tristeza y miedo—, aunque la interacción entre grupo y congruencia alcanzó significación estadística únicamente para alegría y tristeza. Para los escenarios de alegría, las personas autistas mostraron un menor contraste entre las expresiones congruentes (M = 3,28 ± 0,09) e incongruentes (M = 1,53 ± 0,08) que los controles (congruentes: M = 3,54 ± 0,05; incongruentes: M = 1,36 ± 0,07) (B = −0,43 ± 0,15, t(66) = 2,87, p = 0,006). Para los escenarios de tristeza se observó el mismo patrón: personas autistas, M = 2,75 (± 0,10) frente a M = 1,97 (± 0,09); controles, M = 3,18 (± 0,07) frente a M = 1,93 (± 0,07) (B = −0,48, EE = 0,13, t(66) = 3,71, p < 0,001). Para los escenarios de miedo, las puntuaciones mostraron una tendencia en la misma dirección —personas autistas: M = 2,76 (± 0,10) en ensayos congruentes y M = 1,97 (± 0,10) en incongruentes; controles: M = 3,03 (± 0,07) y M = 1,99 (± 0,07), respectivamente—, aunque la interacción no alcanzó significación estadística (B = −0,25, EE = 0,15, t(66) = 1,61, p = 0,112). Al analizar específicamente los ensayos congruentes, las personas autistas otorgaron puntuaciones de predictibilidad significativamente inferiores a las de los controles en las tres emociones: alegría (B = 0,26 ± 0,10, t(66) = 2,46, p = 0,016), tristeza (B = 0,44 ± 0,13, t(66) = 3,49, p < 0,001); y miedo (B = 0,27 ± 0,12, t(66) = 2,25, p = 0,028). Por el contrario, en los ensayos incongruentes no se encontraron diferencias significativas entre grupos para ninguna de las tres emociones: alegría (B = −0,17 ± 0,10, t(66) = 1,64, p = 0,105), tristeza (B = −0,04, ± 0,11, t(66) = 0,36, p = 0,722), y miedo (B = 0,02 ± 0,12, t(66) = 0,17, p = 0,869).

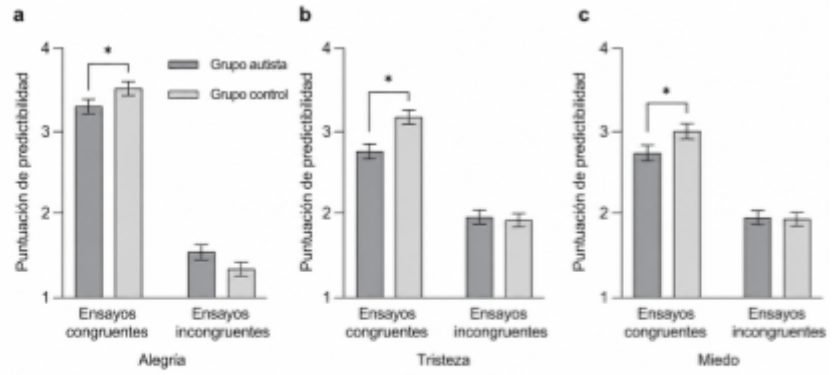


Figura 4. Puntuaciones de predictibilidad de los individuos autistas y del grupo control según la emoción evocada por el escenario y la condición de correspondencia. Se muestran las medias y los errores estándar de los grupos autista (gris oscuro) y control (gris claro). El efecto de correspondencia (*match effect*; diferencia entre las puntuaciones de los ensayos con correspondencia y sin correspondencia) fue menor en los individuos autistas que en los individuos del grupo control. Este efecto se debió principalmente a las diferencias en las puntuaciones de predictibilidad en los ensayos con correspondencia. Los asteriscos indican $p < 0,05$. Extraído de Robles et al., 2024.

En conjunto, estos resultados indican que las diferencias en la estereotipicidad de las predicciones no estuvieron restringidas a una única emoción, sino que mostraron un patrón consistente a través de las distintas categorías emocionales examinadas.

Relación entre las características autistas y la estereotipicidad de las predicciones

El análisis dimensional realizado en la muestra ampliada de participantes sin diagnóstico de autismo (N = 41) mostró que las puntuaciones más elevadas en el AQ-Short se asociaron significativamente con un menor match effect ($R^2 = 0,109$, $F(1, 39) = 4,767$, $t = -2,183$, $p = 0,035$). Por tanto, la relación entre las características autistas y unas predicciones faciales menos estereotípicas no se observó únicamente al comparar categóricamente personas autistas y controles. También dentro de las personas sin diagnóstico de autismo, una mayor presencia de características autistas se relacionó con una menor diferencia entre la predictibilidad atribuida a expresiones congruentes e incongruentes.

Influencia de las predicciones sobre la evaluación social

Finalmente, se examinó si la correspondencia entre la expresión facial observada y la predicción individual del participante influía sobre la posterior evaluación social de la persona observada. Se encontró una relación positiva y estadísticamente significativa entre predictibilidad y agradabilidad en ambos grupos. En el grupo control, las expresiones consideradas más predecibles se asociaron con mayores puntuaciones de agradabilidad ($B = 0,23 \pm 0,03$, $t(66) = 6,70$, $p < 0,001$). La misma asociación se observó en las personas autistas ($B = 0,25 \pm 0,04$, $t(66) = 6,50$, $p < 0,001$). Crucialmente, la interacción entre predictibilidad y grupo no fue significativa ($B = -0,01 \pm 0,05$, $t(66) = 0,25$,

p = 0,801). Por tanto, aunque las personas autistas generaron predicciones sobre las expresiones faciales menos estereotípicas que los controles, utilizaron sus propias predicciones de manera comparable al realizar evaluaciones sociales: en ambos grupos, cuanto mayor era la correspondencia entre la expresión observada y la expresión que el propio participante había esperado encontrar, más favorable era la evaluación de agradabilidad de la persona observada.

Estudio 2. Estudio piloto mediante realidad virtual

Resultados conductuales entre grupos

Los análisis mostraron diferencias significativas entre las personas autistas y los controles en la distribución de la mirada durante la interacción social con el agente virtual. En concreto, las personas autistas presentaron un menor tiempo medio de fijación sobre la región de los ojos, $t(10) = -3,64$, $p = 0,005$, $d = -2,10$, IC 95 % [−3,52, −0,62], acompañado de un mayor tiempo de fijación sobre el fondo de la escena, $t(10) = 4,32$, $p = 0,002$, $d = 2,49$, IC 95 % [0,90, 4,03]. Por el contrario, no se encontraron diferencias significativas entre grupos en el tiempo de fijación sobre la región de la boca ($p > 0,05$). También se examinaron los desplazamientos del punto de mirada en el espacio tridimensional relativo a la cabeza. Las diferencias entre personas autistas y controles no alcanzaron significación estadística en ninguno de los tres ejes: eje X, $t(10) = 0,46$, $p = 0,657$, $d = 0,27$; eje Y, $t(10) = 1,71$, $p = 0,118$, $d = 0,99$; y eje Z, $t(10) = 1,13$, $p = 0,286$, $d = 0,65$. Del mismo modo, no se observaron diferencias significativas entre grupos en la rotación de la cabeza en ninguno de los tres ejes: X, $t(10) = 1,08$, $p = 0,305$, $d = 0,63$; Y, $t(10) = 0,76$, $p = 0,468$, $d = 0,44$; y Z, $t(10) = 0,17$, $p = 0,871$, $d = 0,10$.

En conjunto, los resultados conductuales del estudio piloto identificaron un patrón diferencial especialmente marcado en la distribución de la atención visual: durante la interacción social estandarizada, las personas autistas mostraron una menor fijación sobre los ojos del agente virtual y una mayor fijación sobre el fondo de la escena.

Clasificación mediante aprendizaje automático

Los análisis mostraron que determinadas combinaciones de características de mirada y movimiento permitían alcanzar una elevada capacidad de clasificación en esta muestra piloto. Las combinaciones con mejor rendimiento integraron información procedente de distintas modalidades conductuales, incluyendo medidas relacionadas con

la fijación de la mirada sobre los ojos y el fondo de la escena, las regiones de cabeza y boca, componentes espaciales del vector de mirada y características derivadas de la posición y rotación de la cabeza y las manos. Estos resultados indicaron que la información relevante para diferenciar entre grupos no se concentraba necesariamente en una única señal, sino que podía emerger de la combinación de diferentes características de comunicación no verbal. En particular, las variables relacionadas con la mirada aparecieron de forma recurrente entre las combinaciones con mayor capacidad discriminativa, en consonancia con las diferencias observadas en los análisis estadísticos. No obstante, debido al reducido tamaño de la muestra, las elevadas tasas de clasificación obtenidas en este estudio deben interpretarse exclusivamente como resultados exploratorios. El estudio piloto permitió establecer la viabilidad de utilizar las señales registradas mediante realidad virtual como variables de entrada para algoritmos de aprendizaje automático, pero no proporcionaba una estimación suficientemente robusta de su capacidad de generalización a nuevos participantes. Por este motivo, la siguiente fase de la investigación consistió en ampliar sustancialmente la muestra y evaluar el procedimiento mediante una estrategia de validación más rigurosa. Esta evaluación se llevó a cabo en el Estudio 3. Para una descripción exhaustiva de los algoritmos, combinaciones de características y resultados de clasificación explorados en el estudio piloto, véase Robles et al., 2022.

Estudio 3. Evaluación en la muestra completa mediante inteligencia artificial

El Estudio 3 evaluó en la muestra completa de 75 participantes si las señales multimodales registradas durante la interacción social en realidad virtual permitían diferenciar entre personas autistas y controles. A diferencia del carácter exploratorio del estudio piloto, el modelo se evaluó mediante un procedimiento de validación cruzada estratificada y anidada de cinco particiones, dirigido a obtener una estimación más robusta de su capacidad de generalización.

Rendimiento del modelo de clasificación

El algoritmo supervisado de *Gradient Boosting* alcanzó una exactitud (accuracy) de 0,80, lo que indica que el 80% de los participantes fueron correctamente clasificados. El modelo obtuvo además una sensibilidad (recall) de 0,95, una especificidad de 0,63, una precisión de 0,77 y una puntuación F1 de 0,84. La matriz de confusión permitió examinar el rendimiento por grupo. De las 40 personas autistas, 37 fueron correctamente clasificadas como autistas, mientras que 3 fueron clasificadas como controles. Entre los

35 participantes control, 21 fueron correctamente clasificados como controles y 14 fueron clasificados como autistas. Estos resultados muestran una capacidad especialmente elevada para identificar a las personas autistas, reflejada en la sensibilidad de 0,95, mientras que la identificación de los controles fue más moderada, en consonancia con la especificidad de 0,63.

Métrica	Rendimiento
Exactitud (accuracy)	0,80
Sensibilidad (recall)	0,95
Especificidad	0,63
Precisión	0,77
F1	0,84

Figura 5. Matriz de confusión del modelo de Gradient Boosting

Contribución de las características a la clasificación

Los análisis mediante valores SHAP permitieron examinar qué características del comportamiento no verbal contribuyen en mayor medida a las predicciones del algoritmo. Los resultados mostraron que las medidas relacionadas con la mirada tuvieron una contribución especialmente relevante, destacando los patrones de fijación sobre los ojos, y la boca del avatar, y el fondo de la escena. Las características relacionadas con la posición de la cabeza también aportaron información relevante al modelo. Asimismo, determinadas variables correspondientes al movimiento de las manos izquierda y derecha contribuyeron a la clasificación. Por tanto, aunque las características de mirada destacaron entre las variables más informativas, la diferenciación entre personas autistas y controles se apoyó en información procedente de distintas modalidades de comunicación no verbal.

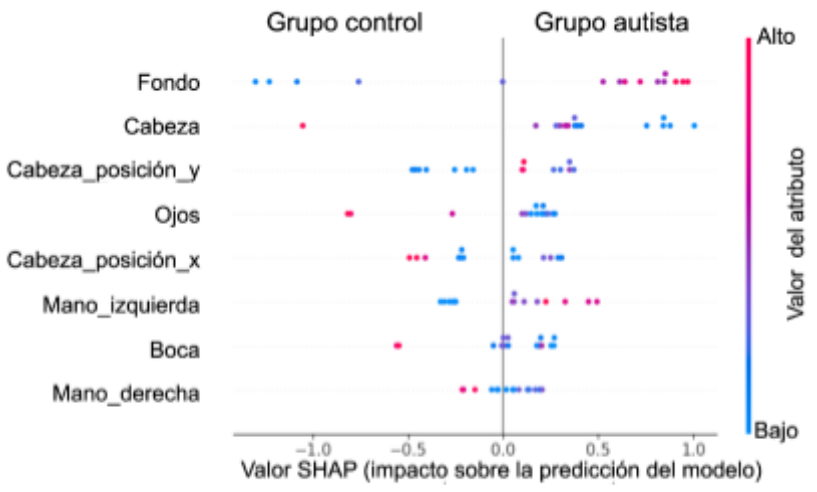


Figura 6. Valores SHAP que muestran el peso de los atributos.

En conjunto, los resultados mostraron que las señales no verbales registradas objetivamente durante una interacción social inmersiva permitieron diferenciar entre personas autistas y controles con una exactitud del 80 % y alcanzando una sensibilidad del 95 %. Estos resultados amplían los hallazgos preliminares del Estudio 2 al demostrar que la aproximación desarrollada en el estudio piloto mantiene capacidad discriminativa al aplicarse a la muestra completa y evaluarse mediante un procedimiento de validación más riguroso.

4. Discusión

El objetivo general de este trabajo fue caracterizar la percepción y producción de señales de comunicación no verbal en personas autistas adultas mediante tres aproximaciones complementarias. En conjunto, los resultados muestran diferencias entre personas autistas y controles tanto en la percepción de señales sociales como en los patrones no verbales producidos durante una interacción. Sin embargo, los hallazgos también revelan procesos compartidos entre ambos grupos y subrayan la importancia de distinguir entre presentar patrones diferentes y utilizar de manera diferente la información social. Además, la progresión desde el estudio piloto hasta su evaluación en una muestra más amplia muestra el potencial de combinar realidad virtual, seguimiento ocular, registro del movimiento e inteligencia artificial para cuantificar objetivamente el comportamiento no verbal.

Las diferencias en comunicación no verbal forman parte de la propia caracterización clínica del autismo (APA, 2022) y, por tanto, su presencia no constituye por sí misma un hallazgo novedoso. La principal aportación de esta investigación reside en cómo se estudiaron estas diferencias. Por una parte, el Estudio 1 examinó las predicciones individuales sobre expresiones faciales sin reducir su evaluación a respuestas consideradas correctas o incorrectas. Por otra, los Estudios 2 y 3 trasladaron el estudio del comportamiento no verbal a una situación social inmersiva, estandarizada y cuantificable, permitiendo registrar simultáneamente señales de mirada y movimiento y analizarlas mediante aprendizaje automático.

Los resultados del Estudio 1 confirmaron la primera hipótesis (H1): las personas autistas mostraron predicciones sobre las expresiones faciales menos estereotípicas que los controles. Aunque ambos grupos consideraron más predecibles las expresiones congruentes que las incongruentes con el contexto emocional, esta diferencia fue significativamente menor en las personas autistas. En particular, las expresiones

estereotípicamente congruentes con el contexto —por ejemplo, una expresión de alegría tras una situación normativamente alegre— coincidieron en menor medida con las expectativas de las personas autistas. Este patrón se observó a través de las diferentes categorías emocionales estudiadas, aunque la interacción entre grupo y congruencia alcanzó significación estadística únicamente para alegría y tristeza. Estos resultados son coherentes con trabajos previos que han descrito diferencias en el procesamiento de expresiones emocionales en el autismo (Harms et al., 2010; Lozier et al., 2014; Uljarevic & Hamilton, 2013; Yeung, 2022), pero permiten plantear una interpretación diferente. En lugar de asumir que una menor coincidencia con una respuesta normativa refleja necesariamente una menor capacidad para reconocer correctamente una emoción, los resultados sugieren que las personas autistas pueden mantener expectativas menos estereotípicas acerca de cómo debería expresarse otra persona en un determinado contexto. Este hallazgo puede interpretarse desde diferentes aproximaciones teóricas. La teoría de la coherencia central débil (Happé et al., 2001) propone una mayor orientación hacia los detalles frente al procesamiento global del contexto, mientras que la teoría de la generalización reducida (Plaisted, 2001) plantea diferencias en la generalización de conocimientos y experiencias. Desde los modelos de procesamiento predictivo, también se ha propuesto que en el autismo podría existir una ponderación diferente de las expectativas previas y los errores de predicción (Cannon et al., 2021; Pellicano & Burr, 2012). Una mayor consideración de características específicas de cada situación podría contribuir a generar modelos internos más detallados y, consecuentemente, predicciones menos generalizables o estereotípicas. No obstante, la tarea empleada no permitió determinar el contenido concreto de las predicciones de los participantes. Por ello, estos resultados no permiten establecer si las predicciones de las personas autistas eran más detalladas, incorporaban aspectos contextuales diferentes o se actualizaban de manera distinta ante nueva información. Tampoco permiten discriminar entre diferentes propuestas dentro de los modelos de procesamiento predictivo, como la hipótesis de priorsatenuados (Pellicano & Burr, 2012) o las propuestas relacionadas con una ponderación inflexible de los errores de predicción en función de la volatilidad contextual (Allenmark et al., 2021; Shi et al., 2024). Estas posibilidades requieren paradigmas específicamente diseñados para manipular las propiedades del contexto y caracterizar directamente el contenido de las expectativas.

Un resultado especialmente relevante fue que la H2 no fue confirmada. A pesar de las diferencias en la estereotipicidad de las predicciones, las personas autistas y los

controles utilizaron sus propias expectativas de forma comparable al evaluar socialmente a otras personas. En ambos grupos, cuanto mayor era la correspondencia entre la expresión observada y la predicción individual, mayor era la agradabilidad atribuida a la persona, y la magnitud de esta relación no difirió significativamente entre grupos. Este resultado introduce una distinción importante entre la naturaleza de una predicción y su utilización funcional. Las personas autistas pueden mantener expectativas que se alejan en mayor medida de las configuraciones normativamente estereotípicas y, al mismo tiempo, utilizar esas expectativas para realizar evaluaciones sociales de manera comparable a los controles. Esta observación es coherente con algunos estudios que no han encontrado diferencias significativas entre personas autistas y controles en determinados juicios sociales (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), aunque otros trabajos sí han descrito diferencias en dimensiones como la confiabilidad, la inteligencia percibida o el atractivo (Adolphs et al., 2001; Couture et al., 2010; Losh et al., 2009; Philip et al., 2010). Una posible explicación de esta heterogeneidad es que numerosos paradigmas previos utilizan criterios normativos de exactitud. Los presentes resultados sugieren que considerar las expectativas propias de cada participante, en lugar de valorar exclusivamente su correspondencia con una respuesta preestablecida, puede aportar información adicional sobre el procesamiento social en el autismo. Esto resulta especialmente relevante en un ámbito caracterizado por una elevada variabilidad individual.

El Estudio 2 abordó la producción de señales no verbales durante una interacción social inmersiva. Los resultados confirmaron parcialmente la H3. Las personas autistas mostraron un menor tiempo de fijación sobre los ojos del agente virtual y una mayor fijación sobre el fondo de la escena que los controles. Estos resultados son coherentes con la literatura que describe una menor orientación hacia la región ocular (Camero et al., 2021; Papagiannopoulou et al., 2014) y una mayor atención hacia elementos del fondo o estímulos no sociales (Kim et al., 2024b; Krishnappa Babu et al., 2023). Sin embargo, no se confirmó la predicción relativa a una mayor fijación sobre la boca. Aunque algunos estudios han descrito este patrón (Camero et al., 2021; Neumann et al., 2006; Riddiford et al., 2022; Wieckowski & White, 2017), otros trabajos, incluida una revisión sistemática y metaanálisis, tampoco han encontrado diferencias significativas entre grupos en la fijación sobre esta región (Papagiannopoulou et al., 2014). Este resultado refuerza la idea de que las diferencias de mirada en el autismo no pueden caracterizarse simplemente como una

redistribución uniforme desde los ojos hacia la boca, sino que dependen de la región considerada, del contexto social y de las características individuales. Las medidas corporales proporcionaron resultados menos consistentes. Las diferencias relacionadas con cabeza y manos fueron más limitadas que las observadas para la mirada, y el registro de las manos estuvo condicionado por la necesidad de utilizar dos mandos durante la experiencia. El propio dispositivo restringía el movimiento espontáneo, por lo que estas medidas deben interpretarse con especial cautela. Esta limitación señala una vía metodológica clara para futuros estudios: el empleo de sistemas de seguimiento de manos sin mandos y visores más ligeros permitiría registrar de forma más natural gestos y movimientos corporales durante la interacción. Más allá de las diferencias concretas entre grupos, una aportación fundamental del Estudio 2 fue demostrar la viabilidad del sistema desarrollado. La RV permitió presentar a todos los participantes una situación social dinámica y estandarizada, manteniendo un elevado control experimental y registrando simultáneamente diferentes señales conductuales. Asimismo, los análisis exploratorios mostraron que estas señales podían utilizarse como entrada para procedimientos de clasificación mediante aprendizaje automático, apoyando la H4. No obstante, el reducido tamaño de la muestra impedía interpretar las elevadas tasas de clasificación exploratorias como una estimación válida de generalización.

El Estudio 3 permitió abordar directamente esta limitación del estudio piloto. Al ampliar el análisis a 75 participantes y emplear un procedimiento de validación cruzada estratificada y anidada, se obtuvo una evaluación más robusta de la capacidad de las señales multimodales para diferenciar entre personas autistas y controles. Los resultados apoyaron la H5. El modelo de Gradient Boosting alcanzó una exactitud de 0,80 y una puntuación F1 de 0,84, con una elevada sensibilidad (0,95) y una especificidad de 0,63. Estos valores resultan comparables a los descritos para herramientas ampliamente utilizadas en la evaluación del autismo, como el ADOS-2, para cuyo Módulo 4 se han comunicado valores de sensibilidad y especificidad en rangos similares. Este resultado resulta especialmente prometedor, considerando que la clasificación se obtuvo exclusivamente a partir de señales conductuales registradas objetivamente durante una interacción social en realidad virtual. Sin embargo, estos resultados no deben interpretarse como evidencia de una herramienta diagnóstica lista para su aplicación clínica. El valor del hallazgo reside en mostrar que patrones objetivos de comportamiento no verbal producidos durante una interacción social controlada contienen información suficiente para permitir una clasificación por encima de lo esperable por azar mediante

técnicas de inteligencia artificial, y que este resultado se mantiene al ampliar sustancialmente la muestra del estudio piloto y aplicar una estrategia de validación más rigurosa. Los análisis SHAP mostraron que las características relacionadas con la mirada —especialmente ojos, boca y fondo— tuvieron una contribución destacada, en línea con el Estudio 2 y con su potencial como marcadores del autismo (Parellada et al., 2022; Wei et al., 2023). La contribución adicional de características de cabeza y manos sugiere que la clasificación se apoya en un patrón multimodal, especialmente relevante para caracterizar la heterogeneidad del autismo más allá de indicadores conductuales aislados.

Limitaciones y futuras líneas de investigación

Los resultados deben interpretarse teniendo en cuenta algunas limitaciones. En el Estudio 1, el paradigma permitió determinar en qué medida las expresiones observadas coincidían con las expectativas de los participantes, pero no conocer el contenido específico de dichas predicciones. Además, se estudiaron únicamente tres emociones y una dimensión de evaluación social. Futuros trabajos deberían ampliar tanto las categorías emocionales como los tipos de evaluación social y desarrollar paradigmas que permitan caracterizar de forma más directa las expectativas individuales.

En los Estudios 2 y 3, el registro de determinadas señales corporales estuvo condicionado por el equipamiento utilizado, especialmente en el caso de las manos, cuyo movimiento natural se encontraba parcialmente restringido por los mandos que los participantes debían de sujetar. El empleo de sistemas de seguimiento corporal y manual sin controladores podría aumentar la validez ecológica del registro y permitir una caracterización más completa del comportamiento no verbal. Aunque el Estudio 3 supone una ampliación sustancial respecto al estudio piloto, el tamaño de la muestra continúa siendo limitado para el desarrollo de modelos de aprendizaje automático con potencial aplicación clínica. Por ello, este estudio forma parte de un proyecto actualmente en curso dirigido a evaluar la robustez, especificidad y capacidad de generalización del sistema. En la actualidad se están analizando datos correspondientes a otros dos grupos clínicos —personas con trastorno por déficit de atención e hiperactividad (TDAH) y con trastorno límite de la personalidad (TLP)— con el objetivo de determinar si los patrones identificados permiten diferenciar específicamente el autismo de otras condiciones clínicas, y no únicamente de participantes control. Asimismo, siguiendo las buenas prácticas metodológicas en aprendizaje automático, está prevista la validación externa del algoritmo en una muestra independiente, ya recogida y actualmente en fase de análisis.

Esta validación permitirá evaluar el rendimiento del modelo en participantes que no han intervenido en su desarrollo y proporcionar una estimación más rigurosa de su capacidad de generalización.

Conclusiones

En conjunto, esta investigación muestra que la comunicación no verbal en personas autistas adultas combina patrones diferenciales y procesos compartidos con personas no autistas. En percepción, las personas autistas mostraron predicciones sobre expresiones faciales menos estereotípicas, pero las utilizaron de forma comparable a los controles en sus evaluaciones sociales, cuestionando interpretaciones basadas exclusivamente en criterios normativos y destacando la relevancia de las expectativas individuales. En producción, la realidad virtual permitió cuantificar señales espontáneas durante una interacción social estandarizada. Las principales diferencias se observaron en la mirada, con menor fijación en los ojos y mayor orientación hacia el fondo. Finalmente, los Estudios 2 y 3 mostraron la viabilidad de emplear estas señales para la clasificación mediante aprendizaje automático. En conjunto, los resultados apoyan un estudio multimodal, objetivo y cuantificable de la comunicación no verbal.

Información adicional sobre la investigación

Los Estudios 1 y 2 han sido publicados previamente en revistas científicas internacionales de alto impacto y sometidos, por tanto, a procesos de revisión por pares. Ambos estudios representaron mi Tesis Doctoral. El Estudio 3 forma parte de una línea de investigación actualmente en curso y presenta resultados todavía no publicados; el manuscrito científico correspondiente, basado en los resultados incluidos en este trabajo, se encuentra actualmente en preparación. Esta línea de investigación se desarrolla en el marco de un proyecto financiado por la Fundación Alemana para la Investigación Científica (DFG), cuya finalización está prevista para finales de 2027. El proyecto incorpora un enfoque de investigación participativa, contando con la participación de personas autistas a lo largo de las distintas fases del proceso de investigación. Todos estudios han sido desarrollados por un equipo de investigación multidisciplinar, cuya composición puede consultarse en las publicaciones correspondientes: Robles, M., Ramos-Grille, I., Hervás, A., Duran-Tauleria, E., Galiano-Landeira, J., Wormwood, J.B., Falter-Wagner, C.M., & Chanes, L. (2024). Reduced stere-otypicality 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

y 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.

5. Referencias

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

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

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

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

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

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. https://doi.org/10.1177/13623613221142383

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? *Jour-nal of Autism and Developmental Disorders*, *53*(2), 677–687. https://doi.org/10.1007/s10803-020-04425-3

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

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

Caulfield, M. D., 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

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

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*, 17008. https://doi.org/10.1038/s41598-018-35215-8

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 fre-quency-specific temporal dynamics. *Cerebral Cortex, 33*(13), 8421–8430.https://doi.org/10.1093/cercor/bhad127

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

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2014). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science, 18*(2), 327–334. https://doi.org/10.1111/desc.12218

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 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., & Mottron, L. (2016). Atypical social judgment and sensitivity to perceptual cues in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 46*, 1574–1581. https://doi.org/10.1007/s10803-014-2208-5

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*(2), 127–138. https://doi.org/10.1038/nrn2787

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

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., 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

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., Posthuma, D., & 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., Xu, F., & Barrett, L. F. (2020). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology, 55*(9), 1830–1849. https://doi.org/10.1037/dev0000686

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*, 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. *Child and Adolescent Psychiatric Clinics of North America, 30*(3), 557–571. https://doi.org/10.1016/j.chc.2020.02.006

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, S. Y., Udhnani, M., & Lecavalier, L. (2024a). Heterogeneity in autism spectrum disorder explained by social-communicative and restricted repetitive behavior balance subgroups. *Research in Autism Spectrum Disorders*, *114*, Article 102387. https://doi.org/10.1016/j.rasd.2024.102387

Kim, N. Y., He, J., Wu, Q., Dai, N., Kohlhoff, K., Turner, J., Paul, L. K., Kennedy, D. P., Adolphs, R., & Navalpakkam, V. (2024b). Smartphone-based gaze estimation for in-home autism re-search. *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

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 atten-tional 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., & 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

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 Spec-trum Disorder. *Journal of Autism and Developmental Disorders*, *49*(11), 4535–4546. https://doi.org/10.1007/s10803-019-04163-1

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., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*(6), 694–701. https://doi.org/10.1001/archpsyc.63.6.694

Lord, C., Luyster, R. J., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) manual: Toddler module*. Western Psychological Services.

Lorenzo, G., Lledó, 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

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-Badia, 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

Maddox, B. B., Brodkin, E. S., Calkins, M. E., Shea, K., Mullan, K., Hostager, J., Mandell, D. S., & Miller, J. S. (2017). 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. https://doi.org/10.1007/s10803-017-3188-z

Maskey, M., Lowry, J., Rodgers, J., McConachie, H., & Parr, J. R. (2014). Reducing specific phobia/fear in young people with autism spectrum disorders through a virtual reality environment intervention. *PLoS ONE, 9*(7), e100374. https://doi.org/10.1371/journal.pone.0100374

Maskey, M., Rodgers, J., Grahame, V., Glod, M., Honey, E., Kinnear, J., Labus, M., Milne, J., Minos, D., McConachie, H., & Parr, J. R. (2019). 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*, 1912–1927. https://doi.org/10.1007/s10803-018-3861-x

Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 35*(6), 655–668. https://doi.org/10.1080/13803395.2013.809700

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

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

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*, 660. 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

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

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*(1–3), 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 J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 149–169). Lawrence Erlbaum Associates.

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.). Sage.

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

Sahin, M., Jones, S. R., Sweeney, J. A., Berry-Kravis, E., Connors, B. W., Ewen, J. B., Hartman, A. L., Levin, A. R., Potter, W. Z., & Mamounas, L. A. (2019). Discovering translational biomarkers in neurodevelopmental disorders. *Nature Reviews Drug Discovery, 18*(4), 235–236. https://doi.org/10.1038/d41573-018-00010-7

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

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 Updat-ing 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 Cog-nition: The Shared-Attention System. *Perspectives on Psychological Science*, *16*(3), 553–576.https://doi.org/10.1177/1745691620953773

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

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

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

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

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

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

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

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., 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., 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., & 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