



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE PSICOLOGÍA**

**DIVISIÓN DEL SISTEMA DE UNIVERSIDAD  
ABIERTA**

**Las Propiedades de los Estímulos Visuales como  
Desencadenantes de la Respuesta a la Atención Conjunta  
en los Trastornos del Espectro Autista. Un Análisis  
Computacional**

**TESIS**

**QUE PARA OBTENER EL TÍTULO DE**

**LICENCIADA EN PSICOLOGÍA**

**P R E S E N T A**

**Diana Nohelí Sinsún Medina**

**ASESOR**

**Dr. Ángel Eugenio Tovar y Romo**

**REVISORA**

**Dra. Ruth Campos García**



**Ciudad Universitaria, CDMX, 2022**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Este trabajo fue apoyado por el Consejo Nacional de Ciencia y Tecnología CONACyT, México. Proyecto “Modelamiento del Desarrollo del Lenguaje en el Síndrome de Down y los Trastornos del Espectro Autista”, CB – 285152.

## AGRADECIMIENTOS

A la *Universidad Nacional Autónoma de México* por las increíbles oportunidades de conocimiento y aprendizaje que me ha brindado y que me han permitido (de)crecer personal y profesionalmente.

A mi incomparable tutor, el *Dr. Ángel Tovar* por invitarme a formar parte del Laboratorio de Desarrollo Cognitivo y del Lenguaje, siendo la primera persona en creer que yo podría realizar un proyecto de investigación. Gracias también por el paciente acompañamiento y constructiva guía en todo este proceso que, sin duda, ha sido un laberinto de tránsito complejo.

A mi queridísima revisora, la *Dra. Ruth Campos* por ser desde aquel primer encuentro un claro referente del camino correcto. También por todos los saberes que con gusto y pasión inmensa me ha transmitido, por la grandiosa oportunidad de formar parte de Traberitea y trabajar con un equipo maravilloso en pro de mi formación como investigadora. Por supuesto, gracias también por todo el tiempo invertido en la revisión de este trabajo, por sus fantásticos comentarios y preguntas que permitieron nuevas reflexiones y horizontes.

A mi inmejorable sínodo, *Dra. Natalia Arias Trejo, Dra. Patricia Romero y Dra. Itzel Galán*. Gracias a cada una por las valiosas y respetuosas aportaciones que hicieron para que mi trabajo alcanzara un mejor nivel.

A mi familia, definitivamente.

A *Yazmín, mi incansable madre*, por el esfuerzo que ha realizado desde siempre para darme una vida llena de oportunidades y por alentarme a tomarlas.

A *Vane, mi adorada hermana*, por todo el apoyo material y emocional que me ha brindado constantemente.

A *Valentina, mi sobrina*, porque incluso sin saberlo ha inspirado en gran parte mi quehacer profesional.

A *Ana Brenda, mi amadita* por su comprensiva, amorosa y paciente compañía en la mayor parte de mi proceso de titulación. Gracias también por la valiosísima contención emocional que me ha

brindado en los momentos de mayor duda existencial y por ayudarme a transitar de nuevo hacia el centro de mi balance y bienestar.

A mis bellísimas amistades del laboratorio.

**Angie y Lore**, por su cálido recibimiento en el equipo, por su inspiradora trayectoria y labor académico-profesional y, por supuesto, por los divertidísimos momentos de goce fuera de la Universidad.

**Marco**, por la invaluable compañía y asesoría durante esta terrorífica pero muy enriquecedora curva de aprendizaje. Agradezco infinitamente todo el conocimiento compartido conmigo, la motivación y las palabras de aliento que lograron apaciguarme en episodios de mayor duda y temor.

**Rebe**, amiga adorada que con su infinito amor y motivación me ayuda a confiar en la vida, en mis decisiones y en mi ser. Incontables y amorosísimas gracias por todo el apoyo, las palabras de aliento, los consejos y esta bella amistad que podemos construir en uno de los más importantes momentos de la vida.

A las amistades que han sido parte de mi red de apoyo.

**Adri Pérez**, por su apoyo desmedido y maternal cuidado durante los primeros años de la licenciatura.

**Tania Polo**, por la bellísima amistad y las hilarantes (pero muy constructivas también) conversaciones sobre nuestros proyectos de tesis.

**Jesus Arias**, amigo del alma y dulce compañía en las experiencias más significativas de la vida hasta ahora.

**Alexia**, guía espiritual y mi norte en todo momento. Psicóloga ejemplar, criatura divina llena de magia, respeto y amor.

**Diana Vázquez**, maestra de vida y claro ejemplo del poder de la confianza en las decisiones y el deseo de crecer por encima de todo miedo a lo desconocido y desafiante.

A **mis peques y sus familias** por la confianza que han depositado en mí para trabajar con ustedes durante estos años. Por todo lo que he aprendido con y para ustedes, por ayudarme a saber

diariamente que esta es la labor que quiero realizar y por permitirme ser parte de su vida y aprendizaje.

*A la vida y al universo* por haberme permitido vivir cada uno de los momentos que han hecho posible encontrarme frente al monitor escribiendo precisamente ahora los agradecimientos en un trabajo que ha tomado suficiente tiempo y esfuerzo para materializarse. Gracias por darme esta oportunidad y no otra, gracias por las personas con las que me ha permitido vincularme y a las que ahora puedo también agradecer.

Gracias profundas, infinitas y amorosas.

## ÍNDICE

1.	INTRODUCCIÓN.....	8
2.	MARCO TEÓRICO .....	11
2.1.	Atención Conjunta.....	11
2.2.	La Respuesta a la Atención Conjunta (RJA) y su Desarrollo Típico.....	14
2.3.	Habilidades de Atención Conjunta en el Trastorno del Espectro del Autismo .....	22
2.4.	Modelos neurocomputacionales y su aplicación a las ciencias cognitivas .....	30
3.	PLANTEAMIENTO DEL PROBLEMA.....	34
4.	JUSTIFICACIÓN Y OBJETIVOS .....	36
5.	HIPÓTESIS .....	38
6.	MÉTODO .....	39
6.1.	Redes Neuronales para Simulaciones de Dirección (DIR) .....	39
6.2.	Redes Neuronales para Simulaciones de Movimiento (MOV).....	40
6.3.	Patrones de Estimulación .....	42
6.5.	Modelamiento del Trastorno del Espectro Autista.....	47
7.	RESULTADOS Y DISCUSIÓN.....	50
7.1.	Desarrollo típico (DT).....	50
7.2.	Trastorno del Espectro del Autismo (TEA) .....	54
7.2.1.	Hiperfocalización (HF) .....	54
7.2.2.	Percepción atípica (PA) .....	58
7.2.3.	Hiperfocalización y percepción atípica (HF+PA).....	59
8.	CONCLUSIONES.....	62
9.	REFERENCIAS .....	66

## **RESUMEN**

La Respuesta a la Atención Conjunta permite coordinar la atención con un/a compañero/a de interacción al seguir pistas sociales indicadas por la mirada, el giro de la cabeza o gestos. Una hipótesis del desarrollo sugiere que esta habilidad podría recaer inicialmente en la percepción del movimiento de la cabeza y se refina hasta responder ante la dirección final de la mirada. En los Trastornos del Espectro Autista se presentan conductas reducidas de atención conjunta, el desarrollo de las cuales es el objetivo de varias intervenciones terapéuticas y podría beneficiarse de un conocimiento preciso sobre la manera en la que las pistas sociales logran desencadenar (o fallan en desencadenar) la Respuesta a la Atención Conjunta. En esta tesis se pone a prueba la hipótesis del desarrollo a través de la creación de un modelo computacional simulando tareas de seguimiento de mirada, y se explora el efecto de las diferencias en la cantidad de información y dinámicas temporales de pistas sociales y no sociales. Dicho modelo es contrastado con estudios empíricos previos y se sugieren trayectorias del desarrollo viables de Respuestas a la Atención Conjunta típicas y atípica proponiendo la relevancia de la cantidad de información percibida por encima de su valor social.

**Palabras clave:** atención conjunta, Trastornos del Espectro Autista, modelo computacional, seguimiento de mirada, pistas sociales y no sociales.

## **ABSTRACT**

The Response to Joint Attention allows coordinating attention with a partner by following social cues as indicated by the gaze, head turn or gestures. One developmental hypothesis suggests that this ability may initially rely on the perception of head motion and refines until responding to the final gaze direction. The Autism Spectrum Disorders present reduced joint attention behaviors,

the development of which is the goal of many therapeutic interventions and could benefit from a precise knowledge of how social cues trigger (or fail to trigger) the Response to Joint Attention.

In this study, we test the developmental hypothesis. We developed a computational model simulating gaze following tasks and explored the effect of differences in the amount of information and temporal dynamics of social and non-social cues. Our model is contrasted with previous empirical studies, and we suggest viable developmental trajectories of typical and atypical Response to Joint Attention, proposing the relevance of the amount of perceived information over its social value.

**Key words:** *joint attention, Autism Spectrum Disorders, computational model, gaze following, social / non-social cues.*

## 1. INTRODUCCIÓN

Cuando apenas me introducía en el mundo de la intervención terapéutica en infantes dentro del Espectro del Autismo (TEA), una compañera me pidió prestar atención a un ejercicio de estimulación de la atención conjunta llamado “Torre no verbal”. El ejercicio consistía en colocarse frente al niño o niña y sostener dos cubos o bloques, uno en cada mano, indicando “con la mirada” el cubo que debía tomar para comenzar o continuar la construcción de una torre de cubos. En ese momento la instrucción parecía sencilla, pero observando con detenimiento, me di cuenta de que varias características cambiaban en la interacción a través del tiempo, y más aún si una terapeuta diferente lo trabajaba. A veces, las terapeutas movían solo los ojos, otras veces acompañaban ese movimiento de una pequeña inclinación de la cabeza, otras tantas movían la mano que sostenía el cubo solicitado e, incluso, podían hacer las tres conductas al mismo tiempo; a veces el ensayo tenía el resultado esperado, que el niño o la niña coordinara la atención para tomar el objeto indicado y utilizarlo en el juego, pero otras no.

Este episodio representa la piedra angular del presente trabajo, nacido de la absoluta incertidumbre del mejor actuar para estimular la atención conjunta. ¿Debía entonces mover solo los ojos?, ¿convenía mover mi cabeza o tal vez agitar mi mano?, ¿qué tal si era mejor agitar la mano y coordinar el movimiento de mi cabeza y de mi mirada hacia el sitio del movimiento? Ciertamente no lo sabía y nadie atinaba a darme una respuesta convincente. Estaba dudosa sobre qué elemento de toda aquella situación permitía o no la manifestación de la atención conjunta: los ojos, la cabeza, la mano; todo ello moviéndose o solo orientándose. Finalmente, deseaba también saber si en realidad existía una única manera de hacerlo o incluso ante toda esa variabilidad, era posible alcanzar el “mismo resultado”.

Fue así como comencé a diseñar una tarea para evaluar el seguimiento de mirada en infantes de manera longitudinal y evaluar las diversas posibilidades informativas: la mirada, la cabeza y un estímulo no social, todas ellas en movimiento o estáticas.

Lo que inicialmente era una tarea diseñada en computadora con fotos de una gran amiga como estímulos visuales, pronto tuvo que ser abstraído en un modelo computacional que, en una situación pandémica en la que el contacto personal era imposible, me permitiera simular el comportamiento de una población con desarrollo típico y una población con TEA frente a situaciones desencadenantes de seguimiento de mirada, caracterizadas por la variabilidad de estímulos respecto a su valor social, la cantidad de información y sus propiedades dinámicas.

El objetivo de este abordaje fue acercarnos a una propuesta coherente la trayectoria del seguimiento de mirada en desarrollo neurotípico y en el espectro del autismo a partir de la evaluación y contraste de teorías innatistas y neuroconstructivistas propuestas anteriormente. De manera simultánea, modelar redes computacionales nos permitió explorar qué modificaciones al modelo típico hacen plausible simular el desarrollo de participantes con TEA, con la intención de responder a las siguientes preguntas: ¿qué diferencias en los mecanismos de procesamiento de información de la atención conjunta pueden explicar mejor la trayectoria del desarrollo del seguimiento de mirada y ¿cuáles son las propiedades de los estímulos que resultan más relevantes en TEA?

En la primera parte del documento se encontrará información fundamental para comprender la importancia de la atención conjunta, la conceptualización del seguimiento de mirada y la investigación que se ha realizado sobre esta conducta en desarrollo típico. Se complementará esta información con las aportaciones que se han hecho desde la investigación

con personas dentro del Espectro del Autismo y cómo es posible simular dicho trabajo gracias al uso de los modelos neurocomputacionales.

Posteriormente, se plantea el problema que motivó esta investigación y su objetivo, así como su justificación y las hipótesis devenidas de ello. El método es descrito abordando la creación de las redes neuronales para la simulación de las propiedades dinámicas de nuestros estímulos, los patrones de estimulación diseñados y las manipulaciones realizadas para el modelamiento del Trastorno del Especto Autista.

Para finalizar, se presentan los resultados analizando el comportamiento de la simulación de Desarrollo Típico y la simulación del Trastorno del Espectro Autista al mismo tiempo que se discute sobre ellos para terminar exponiendo nuestras conclusiones, las limitantes y las implicaciones y posibilidades futuras de nuestro trabajo.

## **2. MARCO TEÓRICO**

### **2.1. Atención Conjunta**

La atención es una función cognitiva compleja de vital importancia en el desarrollo humano que consiste en focalizar selectivamente nuestra consciencia, filtrando y desechando información no relevante (James, 1890; Luria, 1984). Mediante diversos mecanismos neuronales, nos permite manejar el constante flujo de la información sensorial y trabaja para lograr: 1) una percepción precisa de los objetos y la ejecución acertada de acciones sobre éstos, especialmente si hay otros objetos o acciones disponibles; 2) aumentar la velocidad de las percepciones y acciones para preparar el sistema que las procesa; y 3) sostener el uso de recursos cognitivos en la percepción o acción el tiempo que sea necesario (Bench et al., 1993; Desimone & Duncan, 1995; Posner & Petersen, 1989)

La capacidad de concentrarse en un objeto determinado puede adoptar diferentes formas, según factores como su amplitud (e.g. reducida-extendida), dirección (e.g. interna-externa), la actitud del individuo (e.g. voluntaria-involuntaria), las manifestaciones motoras y fisiológicas (e.g. abiertas-encubiertas) y la modalidad sensorial en uso (e.g. auditiva-visual) (Styles, 2010).

Una situación muy sencilla y común que nos ilustra un momento de atención es cuando nos encontramos frente a la televisión viendo nuestro contenido favorito con la voluntad y deseo absolutos de ver los detalles de cada personaje, de escuchar los diálogos, de comprender la trama, de predecir su curso, entre otras cosas; todo esto, ignorando otros estímulos como la música vecina, el sonido de los vegetales asándose y comenzando a quemarse, la palomilla que vuela cerca del foco buscando calor, el rechinar del sillón al movernos y un sinnúmero más de estímulos. En ese momento somos solo nosotros/as atendiendo de manera focalizada a lo que transmite la televisión: un ser humano y un foco atencional.

Si bien un proceso atencional es de naturaleza individual, nuestra condición de seres sociales nos permite un tipo más complejo de atención cuando logramos acoplarla con la de otra persona, experimentando episodios de **atención conjunta**.

La atención conjunta es la habilidad que permite coordinar la atención de una persona con la de otra de manera triádica. Es decir, con la participación del ser mismo, otra persona y un tercer objeto, evento o símbolo (Mundy, 2016). El desarrollo de la atención conjunta representa un proceso particularmente esencial, cumpliendo con una función muy importante durante el periodo prelingüístico al permitir la comunicación de información básica de objetos de interés o incluso de deseos (Butterworth & Jarrett, 1991) entre bebés y las personas que le cuidan. De esta manera, la atención conjunta brinda las bases para la vivencia compartida de experiencias, lo que resulta de gran importancia en los procesos de interacción social, comunicación, adquisición del lenguaje y de la capacidad referencial, por lo que adopta asimismo un papel fundamental en el desempeño del área social (Carpenter et al., 1998).

Retomando la situación de la televisión, pensemos que mamá llega del trabajo y nos observa sentados/as en el sillón mirando atentamente algo, de tal manera que, siguiendo el rastro de nuestra mirada, llega a la televisión, donde encuentra nuestro programa favorito, nos giramos hacia ella al saber su presencia, cruzamos la mirada y ambas personas sonreímos. Ella ya sabe lo que estamos haciendo y nosotros/as sabemos que lo sabe; si ocurriera algo relevante para ambas en la televisión, sería sencillo hacer algún comentario y saber de lo que estamos hablando.

En el episodio descrito, la madre realiza dos acciones importantes: por una parte, ella se **orienta** hacia nuestro punto focal, es decir, dirige sus órganos sensoriales, ya sea de manera voluntaria o refleja, hacia la televisión; además de ello, **procesa** la información, infiriendo que

ese objeto es el que estamos atendiendo. La atención conjunta, por lo tanto, no consiste solo en orientarse hacia un estímulo, sino en procesar adicionalmente la información que este nos brinda.

La atención visual conjunta (JVA, por las siglas en inglés de *Joint Visual Attention*) es la modalidad más común de la atención conjunta y ha sido mayormente definida como la habilidad para seguir la dirección de la mirada de otra persona o, simplemente “mirar hacia donde alguien más está mirando” (Butterworth & Jarrett, 1991). Sin embargo, el término “mirar” resulta ambiguo, en tanto que podría referirse a solo voltear a ver algo o atenderlo y centrar nuestra atención en él. Al respecto, la distinción entre “orientación” y “atención” cobra importancia, pues la investigación que se ha hecho en el campo ha considerado de manera casi equitativa: 1) la orientación refleja (Gregory et al., 2016; Kylliäinen & Hietanen, 2004; Senju et al., 2004; Swettenham et al., 2003; Valenza et al., 2015), 2) los cambios atencionales reflejos (Farroni et al., 2000; Hood et al., 1998; Vecera & Rizzo, 2006) y 3) el seguimiento de mirada (Brooks & Meltzoff, 2005; Del Bianco et al., 2018; Michel et al., 2021).

Las 3 conductas mencionadas en el párrafo anterior forman parte de lo que se ha nombrado *Conductas de Atención Conjunta*, las cuales son fundamentales en el mundo social, ya que se encuentran implicadas en el mecanismo básico de desarrollo humano: la interacción. Mundy (2016) propuso la **Iniciación de la Atención Conjunta** (IJA, siglas en inglés de *Initiation of Joint Attention*) y la **Respuesta a la Atención Conjunta** (RJA, siglas en inglés de *Response to Joint Attention*) como conductas que cumplen con la función de compartir la atención con un/a compañero/a de interacción o monitorear la de la misma, ya sea iniciando conductas de atención social de manera espontánea, como establecer contacto visual o alcanzar/señalar/dar un objeto (IJA) o siguiendo pistas sociales como la dirección indicada por la

mirada, un giro de la cabeza y/o un gesto realizado por otra persona (RJA). En esta tesis el foco principal es sobre la RJA, con lo cual se detalla más de esta en la siguiente sección.

## **2.2. La Respuesta a la Atención Conjunta (RJA) y su Desarrollo Típico**

La atención a las pistas sociales que recibimos de quienes nos rodean, que incluyen: expresión facial, gestos, orientación del cuerpo, lenguaje y, mayormente, la mirada, es una capacidad fundamental para el desarrollo de diversas habilidades. Gracias a ella, se establece la oportunidad de interactuar en el mundo y aprender a través del modelado o imitación para alcanzar los hitos importantes del desarrollo (Wang et al., 2020).

Muy temprano en la vida, los seres humanos que se desarrollan de manera típica comienzan a manifestar conductas de atención social, lo que implica que son capaces de atender selectivamente a las pistas sociales de su entorno, de tal manera que, al nacer, muestran una atención preferencial hacia voces familiares (DeCasper & Fifer, 1980), hacia el movimiento biológico (Shiffrar et al., 1997; Simion et al., 2008), hacia estímulos tipo rostro (Johnson, 2005) y hacia rostros (Haith et al., 1977).

La atención preferencial hacia los estímulos mencionados posee gran relevancia, siendo particularmente la tendencia a mirar hacia los rostros la que facilita el procesamiento de estos (Diamond & Carey, 1986; Tanaka & Gauthier, 1997) y conduce, a su vez, a dirigir la atención hacia la región de los ojos (Haith et al., 1977). La mirada, como un elemento central que desde muy temprana edad se identifica como una manera eficaz de sintonizarse con el/la otro/a (Tomasello et al., 2007) se convierte en un recurso informativo importante al indicar fuentes de interés o peligro en el ambiente, por lo que atenderla con frecuencia favorece la comunicación a través de la atención conjunta.

En situaciones naturales, la mayoría de los episodios de atención conjunta surgen como resultado del monitoreo que la madre (como cuidadora primaria) realiza de la línea de mirada de los/las infantes/as (Schaffer, 1984), no obstante, en la infancia temprana también se es capaz de manifestar RJA's de manera contingente a los cambios en el foco de la atención visual de la madre o de otra persona. En otras palabras, la capacidad de atención conjunta es un fenómeno recíproco en el que tanto una persona adulta como un/a bebé pueden estar alerta respecto a un objetivo espacial de los cambios de mirada al realizar el seguimiento de esta.

Gracias al interés que la atención conjunta generó en el campo de la investigación y teniendo como base los avances y descubrimientos en el área de la percepción visual (Fantz, 1961), varios/as investigadores/as decidieron comenzar una larga, y muy interesante, carrera en el descubrimiento del desarrollo de la atención conjunta y el seguimiento de la mirada desde el nacimiento, siendo Scaife y Bruner (1975) los primeros psicólogos en crear un metodología para el estudio de la actual Respuesta a la Atención Conjunta, dando lugar a lo que se conoce como el prototípico **Paradigma de Seguimiento de la Mirada o Paradigma de Atención Conjunta**.

En este paradigma, un/a investigador/a establece una interacción cara a cara con cada infante. Después de establecer contacto visual, el/la investigador/a gira su cabeza de manera silenciosa hacia un lado para mirar hacia un estímulo (una ligera luz en el experimento original) brindando una pista al/a la infante/a para cambiar la dirección de su atención hacia la de él/ella. Se llevan a cabo ensayos de manera subsecuente en los que el/la experimentador/a reestablece primero el contacto visual con el/la infante/a y después brinda una pista para generar un cambio atencional. Los casos en los que el/la infante/a cambia la dirección de su mirada para alinearla con la del/de la investigador/a son considerados episodios de atención (visual) conjunta.

Scaife y Bruner (1975) llevaron a cabo su investigación con 34 bebés de entre 2-14 meses, quienes participaron en una sesión experimental de solo 2 ensayos, consistiendo cada uno de ellos en un cambio en la mirada del/de la experimentador/a hacia cada lado. En cada ensayo, el/la experimentador/a giraba su cabeza y ojos de manera conjunta  $90^\circ$  para fijar la mirada en un estímulo que no era visible para el/la infante/a y consideraban un establecimiento exitoso de atención (visual) conjunta si los/las infantes/as miraban en la misma dirección que el/la experimentador/a en por lo menos uno de los dos ensayos. Sus resultados mostraron que el 30% de infantes ya establecían atención (visual) conjunta a los 2 meses, cifra que aumentaba de manera consistente en relación con la edad, de tal manera que a los 11-14 meses, el 100% de infantes que participaron en la investigación demostraron girar su cabeza en la dirección correcta en por lo menos uno de los dos ensayos experimentales. De acuerdo con sus resultados, los investigadores propusieron que el seguimiento de la mirada se desarrolla a partir de una sensibilidad específica hacia estímulos sociales o bien como una respuesta básica a señales perceptuales como las que algunos animales no humanos usan para comunicar peligro.

Aunque los fundamentos metodológicos y la propuesta de desarrollo de Scaife y Bruner (1975) generaron algunas críticas (Corkum & Moore, 1998), su trabajo y los juicios alrededor de él establecieron las bases para los estudios subsecuentes, en los que los/las investigadores/as se propusieron estudiar el fenómeno de la atención conjunta infantil con mayor detalle para distinguir entre los muchos mecanismos y causas que podrían estar implicados en la reorientación de la atención de los/las infantes/as.

Butterworth, en sus trabajos con Cochran (1980) y Jarrett (1991) propuso que el seguimiento de la mirada (que también llamó *coorientación*) se desarrolla a partir de una predisposición innata, por lo que cualquier restricción (*constraint*) en la maduración explicaría

los cambios en el patrón de desarrollo durante el primer año de vida, mientras los factores ambientales tendrían una influencia mínima en él.

En tono similar a posturas innatistas, Baron-Cohen (1995a, 1995b) propuso la existencia del **Detector de Dirección Ocular** (EDD, por las siglas en inglés de *Eye Direction Detector*); un sistema neurocognitivo especializado en la detección rápida de la dirección a la que miran los ojos de otro organismo, cuyo origen filogenético evolutivo se explicaría a partir de su considerable importancia adaptativa.

De acuerdo con lo postulado, el EDD se activaría en respuesta a estímulos “tipo ojos”, teniendo una activación más fuerte cuando la orientación de esos estímulos es hacia uno/a mismo/a y favoreciendo cambios atencionales de los/as infantes/as incluso antes de la emergencia formal de las conductas de atención conjunta.

Esta teoría encontró su primer apoyo empírico en las investigaciones de Hood y colaboradores (1998), quienes llevaron a cabo dos experimentos para examinar si los/as infantes/as son capaces de manifestar cambios en su atención visual ante los cambios en la mirada de los/as adultos/as, particularmente el giro de los ojos. Ambos experimentos tuvieron como base el paradigma de seguimiento de mirada, aunque con una situación más controlada a partir de la computarización de la tarea. En el primer experimento, el/la infante/a era sentado/a en el regazo de su cuidador/a frente a un monitor en el que aparecía el rostro de una mujer parpadeando hasta que el/la infante/a centraba su atención en él. Posteriormente, aparecía la imagen de la misma mujer con los ojos orientados hacia la derecha o hacia la izquierda y, ante la desaparición de ella, se mostraba en el monitor un estímulo *target* que aparecía hacia el mismo lado en el que se habían orientado los ojos (condición congruente) o hacia el lado opuesto (condición incongruente). En el segundo experimento se mantuvieron las mismas condiciones

del primero, pero el rostro permanecía en la pantalla durante la aparición del estímulo *target*, esto con el propósito de evaluar el comportamiento infantil frente a una situación más ecológica en la que debe ocurrir un desenganche de la atención.

Los/las participantes completaron un promedio de 14.5 ensayos válidos (DE = 3.4, rango 9-20), orientándose hacia el estímulo *target* en el 82% de ellos (DE = 10.5, rango 67-100). Analizando el tiempo transcurrido entre la aparición del estímulo *target* y su orientación hacia él (i.e., latencia) y el acierto-errore al orientarse a la dirección congruente, los investigadores concluyeron que los/las bebés de 3 meses son capaces de detectar la dirección de la mirada cuando son solo los ojos los que la indican, siendo esta detección la que a su vez influye en su propia dirección de la atención. Los resultados sugirieron también que la capacidad temprana e desenganchar y alternar la atención en dirección a la orientación percibida de la mirada puede ser difícil de manifestar en paradigmas convencionales y naturalistas por lo que, aunque la desaparición del estímulo desencadenante es un evento antinatural, su ocurrencia puede ser crítica para que infantes muestren su capacidad en este tipo de tareas.

Mientras las perspectivas innatistas parecían ir ganando territorio en la comprensión del seguimiento de mirada como RJA, una serie de estudios comenzó a ser publicada por Moore y su equipo de trabajo (Corkum & Moore, 1995, 1998; Moore et al., 1997), en la cual se propuso el desarrollo del seguimiento de mirada como un aprendizaje por reforzamiento que implica: 1) factores ambientales (e.g. ubicación espacial, relevancia social y dinamismo de los estímulos) y 2) precursores ontogenéticos relacionados particularmente con el procesamiento de la información ambiental disponible.

Estas nuevas aportaciones establecieron la oportunidad de analizar detalladamente el rostro y la mirada como dos de las pistas sociales más relevantes para la RJA, influyendo de

maneras particulares en su emergencia, manifestación y desarrollo con base también en sus propiedades dinámicas.

Aunque los análisis por tipo de estímulos y propiedades dinámicas ya se vislumbraban en los trabajos de Lempers (1979) y representaron una línea de trabajo importante en los años 80 y 90, la propuesta de Baron-Cohen respecto al EDD (1995a, 1995b) causó un impacto tal que, a inicios del presente milenio, los títulos de diversos artículos científicos comenzaron a abordar el debate respecto a los estímulos iniciadores de atención conjunta más eficaces, considerando el rostro y la mirada como elementos aislados y/o coordinados capaces de transmitir información a través de su movimiento y/u orientación.

Un par de años después de que el “primer soporte inequívoco para la hipótesis del EDD” (Hood et al., 1998) viera la luz, el equipo de Farroni (2000) realizó una serie de experimentos con el propósito de evaluar la importancia del movimiento percibido en la RJA, con la idea de que el diseño experimental de Hood y colaboradores no aislaba la dirección (orientación) de los ojos de su movimiento, por lo que los resultados podrían no haber sido necesariamente producto de una detección innata de la dirección de la mirada, sino de la percepción del movimiento.

De esta manera, tres experimentos fueron conducidos con bebés de 4 a 5 meses para explorar el papel del movimiento a través del mismo paradigma de seguimiento de mirada (*eye-gaze cueing*) de Hood, Willen y Driver (1998). El primer experimento logró replicar los hallazgos previos respecto a que los/as infantes/as manifiestan algunos movimientos anticipatorios que les permiten realizar sácadas (i.e., movimientos oculares rápidos) hacia *targets* periféricos cuando la dirección de los ojos de un rostro localizado al centro de su campo de visión lo indican (Johnson et al., 1991). El segundo experimento reveló que los cambios atencionales de infantes son, en realidad, susceptibles a la dirección del movimiento y no a la

orientación estática de la mirada cuando se modifica el paradigma de tal manera que los ojos se mantienen inmóviles mirando al frente mientras el rostro se desplaza hacia un lado. Finalmente, ambos resultados fueron confirmados por un tercer experimento en el que las señales informativas no lograron causar un cambio atencional bajo condiciones similares a las del diseño del primer experimento cuando no hubo movimiento aparente de las pupilas.

Si bien los descubrimientos hechos hasta ese momento tuvieron gran valor, los aparentes éxitos de los diseños experimentales también parecieron nublar la visión de los/as investigadores/as, pues las evidentes diferencias metodológicas sugieren un interés en apoyar una u otra hipótesis, sin considerar diseños experimentales más complejos que incluyeran explicaciones complementarias, o la posibilidad de que los resultados pudieran corresponder con etapas específicas en el desarrollo de la RJA, siendo que varias ejecuciones conductuales en respuesta a distintos componentes ambientales serían esperados a través del tiempo.

También es importante notar que la mayoría de los estudios enfocados en procesos cognitivos, sin ser los de seguimiento de la mirada la excepción, se suelen centrar en un grupo de datos proveniente de dominios simples, es decir, de áreas del desarrollo aisladas. No obstante, sin resultados provenientes de evaluaciones en diversos dominios de los/as mismos/as infantes de manera longitudinal, no es posible asegurar el momento del desarrollo en el que ocurren cambios cognitivos y se manifiestan algunas habilidades (Karmiloff-Smith et al., 2010).

Notablemente, de manera simultánea al trabajo de los/as investigadores/as ya citados/as, también se llevaron a cabo algunos estudios longitudinales que permitieron el planteamiento de diversos mecanismos presentes con características particulares capaces de favorecer el seguimiento de la mirada en distintas edades y dependiendo de habilidades más generales como la preferencia visual y la percepción del movimiento y del espacio.

Corkum y Moore (1995) examinaron, de manera sistemática, los cambios en las pistas sociales y perceptuales que influyen en el establecimiento infantil de atención conjunta a través de dos estudios realizados con bebés. En el primer estudio (N=60, 6-19 meses) se empleó el paradigma de atención conjunta para indagar en el comienzo de la atención conjunta y las pistas perceptuales que se usan a diferentes edades para determinar la dirección de la mirada de otra persona; este estudio fue complementado por un segundo estudio (N=63, 6-11 meses) donde se abordaron los mismos objetivos, enfatizado particularmente en los orígenes de la Respuesta a la Atención Conjunta.

Los resultados mostraron que los/as bebés son capaces de coordinarse con la dirección de la mirada de otra persona alrededor de los 10 meses, siendo posible demostrarlo de manera confiable al final del primer año de vida. Si bien Corkum y Moore plantearon un comienzo más tardío de la atención conjunta respecto al propuesto por Scaife y Bruner (1975), contaron también con mayor confianza en sus hallazgos al encargarse de las lagunas metodológicas que habían sido criticadas en el estudio de dichos investigadores.

En adición a lo anterior, a partir de este par de experimentos se propuso que, alrededor de los 12 meses de edad, los seres humanos parecen involucrarse en episodios de atención conjunta basándose principalmente en la información relativa a la orientación de la cabeza, pues aunque ya son capaces de detectar cambios en la orientación de los ojos de otra persona (Mayer & Dobson, 1982) y pasan más tiempo escaneando la región de los ojos que ninguna otra área (Hainline, 1978; Haith et al., 1977), el giro de la cabeza representa un movimiento transitorio fuerte y grueso que brinda información más efectiva (Anstis et al., 1969) a esa edad; es decir, es un movimiento más notable. Por otro lado, hasta los 18 meses los/as infantes/as parecen apreciar la importancia de la orientación visual de la otra persona al generar ocasionalmente cambios

atencionales, poseyendo mayor fuerza y probabilidad de dar origen a un episodio de atención conjunta cuando se acompaña de la cabeza, dado que los movimientos combinados señalan de una manera más clara los cambios en el foco atencional de los/as adultos/as que solo el movimiento de los ojos (Butterworth & Jarrett, 1991).

Enmarcados en el paradigma neuroconstructivista, los resultados que se obtuvieron de los estudios longitudinales ya citados representaron un soporte importante para la **Hipótesis del Estrechamiento Perceptual de la Emergencia del Seguimiento de Mirada** (*Perceptual Narrowing Account of GF Emergence*) (Del Bianco et al., 2018), en la que se sugiere un refinamiento en el seguimiento de la mirada, concebido como un proceso de estrechamiento perceptual en el que los/as infantes manifiestan inicialmente una calibración progresiva hacia la presentación de estímulos más familiares (e.g. seguir el movimiento en la forma de la cabeza, luego en compañía de la mirada, para finalmente orientarse preferencialmente hacia donde lo indican los ojos, incluso si la cabeza se encuentra orientada en otra dirección).

Dada la importancia de la Respuesta a la Atención Conjunta, concibiéndola como una habilidad de inicio temprano bajo un proceso de especialización constante en las trayectorias típicas del desarrollo, y aunado al interés teórico por clarificar su manifestación y desarrollo, resulta de gran importancia para este trabajo el indagar en las trayectorias atípicas del desarrollo que presentan patrones diferentes de atención hacia los estímulos IJA y sus características, siendo uno de los ejemplos más notables el que se observa en el Trastorno del Espectro del Autismo.

### **2.3. Habilidades de Atención Conjunta en el Trastorno del Espectro del Autismo**

El Trastorno del Espectro del Autismo (TEA) es una condición del neurodesarrollo que se manifiesta a través de un amplio espectro de alteraciones persistentes en las habilidades de

comunicación social recíproca y en la interacción social, acompañadas de patrones repetitivos y/o restringidos de conducta, actividades e intereses, así como de respuestas sensoriales atípicas que impactan en el desempeño global de las personas que lo viven (American Psychiatric Association, 2014).

Desde su primera descripción en el siglo XX, el espectro autista y su caracterización han motivado la investigación en psicología y áreas afines para identificar las dimensiones implicadas en el TEA. Actualmente, el fenotipo heterogéneo del TEA se describe como el producto de una interacción compleja de grados variables de desarrollo atípico a través de múltiples dimensiones genéticas y bioconductuales (Happé et al., 2006), con impacto en sistemas de desarrollo como el perceptual, motor, social y el atencional (Elsabbagh & Johnson, 2016).

Una de las dimensiones con gran evidencia respecto a su validez y confiabilidad en el diagnóstico e identificación temprana del TEA es la atencional, manifestando atención social y conjunta de manera reducida (Osterling & Dawson, 1994) en comparación con las trayectorias típicas del desarrollo, lo cual influye de manera importante en la adquisición y desarrollo del lenguaje, de la conducta cooperativa y de la cognición social, así como en el procesamiento de información relevante en interacciones sociales (Mundy, 2016).

Wang y colaboradoras (2015) propusieron que los patrones atípicos de atención representan el origen de las dificultades sociales que caracterizan el TEA, siendo la ausencia de habilidades de iniciación y respuesta a la atención conjunta (IJA y RJA) indicadores importantes para el diagnóstico temprano en evaluaciones estandarizadas como la Escala de Observación para el Diagnóstico del Autismo (ADOS, por las siglas en inglés de *Autism Diagnostic Observation Schedule*) (Lord, 2015) y la Escala de Comunicación Social Temprana (ESCS, por las siglas en inglés de *Early Social Communication Scales*) (Mundy et al., 2003). Esta

caracterización es relevante, pues en un marco de conducta atípica acentúa el componente de desarrollo, que va desde habilidades atencionales básicas a habilidades sociales complejas, en lugar de observar el TEA como una condición con dificultades específicas en el componente social.

Las dificultades en el desarrollo de habilidades de atención conjunta observadas en personas con TEA han sido asociadas con los patrones atípicos de atención que manifiestan desde muy temprana edad con respecto a estímulos usualmente considerados como socialmente saliente o relevantes (Astor et al., 2021; Haith et al., 1977; Vlamings et al., 2005; Wang et al., 2020).

Aunque al nacer los seres humanos se orientan y atienden visualmente de manera preferencial hacia el área de los ojos de las personas que les rodean, a partir de los 2 años de edad, los/as infantes/as que más tarde reciben diagnóstico de TEA comienzan a presentar un declive en la fijación hacia los ojos, un patrón que no se observa en infantes/as que no desarrollan el Trastorno del Espectro del Autismo (Jones & Klin, 2013).

A nivel neuroanatómico, varios estudios han reportado anomalías anatómicas y funcionales en el Surco Temporal Superior (STS) en TEA (Redcay, 2008; Zilbovicius et al., 2006), la misma zona en la que células de monos responden selectivamente a la dirección de la mirada percibida (Perrett & Mistlin, 1990). Se ha reportado que esta zona también se activa en estudios de neuroimagen de personas con desarrollo típico (Haxby et al., 2000; Puce et al., 1995) en compañía de la amígdala (Kawashima et al., 1999) cuando se percibe la mirada, por lo que se ha propuesto como un área cerebral específica para su procesamiento.

Dado que la habilidad para seguir la atención de otra persona a través de la dirección de la mirada goza de relevancia filogenética y podría recaer de manera importante en la función del

STS, en las últimas décadas se ha ampliado la investigación en desarrollo típico para comprender cuáles son las pistas que también podrían generar cambios atencionales en TEA.

Algunos estudios han mostrado que en TEA la atención visual es guiada por el movimiento percibido de los ojos, como reportaron Chawarska y su equipo (2003) al realizar un estudio con 12 niños y 3 niñas dentro del espectro autista con edad promedio de 2 años a través del *Visual Attention Cueing Paradigm*. En su estudio contrastaron el movimiento de los ojos en la imagen de un rostro normal y de uno pixelado (considerado no biológico) como señales desencadenantes del seguimiento espontáneo de la mirada utilizando una cámara sensible a la luz infrarroja para grabar la conducta visual de los participantes y analizar la latencia en las sácadas ante ambos estímulos. Después de analizar las grabaciones y evaluar los tiempos de respuesta, los investigadores descubrieron que los/as participantes dentro del espectro del autismo no solo manifestaban una respuesta atencional refleja ante el movimiento de los ojos, sino incluso en un tiempo menor que el grupo con desarrollo típico. Los resultados favorecieron la idea de que la atención visual en TEA puede ser guiada por el movimiento percibido de los ojos, pero también reconocieron que en situaciones naturalistas en realidad no se observa un seguimiento de la mirada, sino una sensibilidad hacia señales direccionales inherentes al movimiento de los ojos.

Llegando a la edad preescolar, los/as niños/as con TEA muestran dificultades atencionales específicas a nivel diádico y triádico, siendo menos responsivos/as que niños/as control con desarrollo demorado sin TEA, según propusieron Leekam y su equipo (2000) a través de una serie de experimentos basados en el paradigma de Corkum y Moore (1998). No obstante, los/as participantes de dichos experimentos no mostraron dificultades para desenganchar la atención y orientarse hacia estímulos target con una mayor velocidad, lo cual

contrasta con los resultados de otros estudios en el área (e. g. Casey et al., 1993; Harris et al., 1999; Wainwright-Sharp & Bryson, 1993).

Ante los primeros resultados en favor de que las personas con TEA responden al movimiento de los ojos y de la cabeza en la infancia, y las posteriores dificultades halladas en preescolares de esta población, el equipo de Swettenham (2003) se dispuso a ahondar en la misma línea de investigación para estudiar si la percepción de los ojos en movimiento podía actuar como desencadenante de la orientación visual refleja en jóvenes con TEA. Llevaron a cabo 2 experimentos con escolares dentro del espectro del autismo de alto funcionamiento (edad promedio de 10 años, 2 meses) usando el paradigma de seguimiento de la mirada, en el que manipularon el tiempo de aparición de la señal en pantalla (100 ms. vs. 800 ms), brindaron la instrucción explícita de “ignorar el movimiento ocular del rostro en la pantalla” e incluso invirtieron el rostro en la pantalla (la pista) para explorar si esta modificación podría tener algún impacto en el procesamiento del rostro dada la modificación de la información configuracional, pudiendo así eliminar la respuesta refleja de seguimiento de mirada. Los resultados llevaron a concluir que hay un componente reflejo en los cambios de atención, ya que se manifiestan incluso cuando se pide no prestar atención al movimiento y la señal aparece en pantalla por muy poco tiempo; adicional y decisivamente, compartieron que el efecto persistía incluso cuando la señal estaba invertida, lo que les llevó a proponer que la dirección del movimiento de los ojos puede representar una poderosa señal para manifestar un cambio atencional en escolares con TEA.

Con un diseño experimental casi idéntico, Kylliäinen y Hietanen (2004) demostraron que la dirección estática de la mirada de otra persona también es capaz de provocar cambios automáticos en la atención visual en escolares con TEA con una edad promedio de 12 años,

quienes en el estudio fueron capaces de orientarse visualmente de igual manera en respuesta a la dirección de la mirada de otra persona o a la indicada por una flecha, sin mostrar una preferencia hacia la señal social.

Con la posibilidad de que ambos: movimiento y dirección de la mirada de otra persona, funcionaran como desencadenantes de cambios atencionales en personas con TEA o en realidad estuvieran siendo confundidos en los diseños experimentales, Rutherford y Krysko (2008) diseñaron un experimento para analizar de manera aislada el movimiento (como una señal no social) y la dirección de los ojos (como una señal social). Para tal propósito, utilizaron un paradigma casi idéntico al de seguimiento de mirada y crearon dos condiciones: en la condición de Movimiento de Ojos, los ojos se movían hacia un target o hacia la dirección opuesta; en la condición de Movimiento de Rostro, la imagen de una cara podía girar hacia la derecha o a la izquierda mientras las pupilas permanecían inertes en la pantalla viendo un punto concreto en la dirección opuesta, lo que consideraron que daba el efecto de oposición entre la dirección final de los ojos y la del movimiento de la cabeza.

Este nuevo diseño fue puesto a prueba con 46 participantes, 23 adultos jóvenes con TEA de alto funcionamiento y 23 con desarrollo típico que registraban sus respuestas con una palanca de mandos (*joystick*). El análisis de los resultados mostró en la condición de Movimiento de Ojos una orientación refleja de los participantes de manera congruente con la dirección final de los ojos en la pista, siendo las personas con desarrollo típico quienes respondían con mayor velocidad. En adición, en la condición de Movimiento de Rostro los participantes continuaron orientándose hacia donde la dirección de la mirada lo había indicado, aunque esta pista fuera opuesta a la del movimiento del rostro, por lo que concluyeron que la dirección final de la

mirada final dirige la atención refleja en ambas poblaciones: TEA y desarrollo típico, de manera más fuerte y consistente que el movimiento del rostro.

La revisión y comparación de los estudios presentados nos permite apreciar diferencias en el efecto de las señales iniciadoras de atención conjunta en distintas etapas del desarrollo y bajo diversas condiciones metodológicas. La variabilidad en los resultados, por lo tanto, parece depender en gran medida de dos aspectos: el diseño experimental de cada estudio y las características de la población participante.

Respecto al diseño experimental, sabemos que la mayor parte de los estudios han utilizado el paradigma clásico de seguimiento de mirada, ya sea de manera tradicional (una persona adulta sentada frente al/a la bebé y orientándose hacia un estímulo target ubicado usualmente a la derecha o izquierda de la habitación) o computarizada (con estímulos presentados en una pantalla). Asimismo, el movimiento se ha manipulado para hacerlo visible u ocultarlo y presentar solo la dirección final del estímulo, se han utilizado estímulos sociales y/o no sociales variando la señal a estudiar e incluso varias conductas han sido registradas como relevantes: la orientación visual refleja, la latencia de las sácadas, la identificación de la ubicación del *target* y/o la anticipación de su aparición. Finalmente, las conductas también se han identificado a través de videos que posteriormente son analizados cuadro por cuadro o con ayuda de rastreadores visuales y de palancas de mandos (*joysticks*).

Adicional a todas las variaciones mencionadas, las diferencias en los resultados a nivel del enganche atencional y tiempo de procesamiento de las señales podrían recaer en diferentes estrategias de procesamiento de la mirada, atribuyéndose comúnmente al desarrollo de las redes cerebrales implicadas en la percepción y la cognición en cada rango de edad estudiado.

De acuerdo con la hipótesis de conectividad, el desarrollo del Trastorno del Espectro del Autismo se caracteriza por un patrón de diferencias significativas en la calidad y cantidad de las características de la red implicada en la transmisión y transferencia de información intra y entre sistemas neuronales locales o distales del cerebro. La *Teoría del Mundo Intenso* sugiere un hiperfuncionamiento de dichos sistemas neuronales, caracterizándose por una hiper-reactividad e hiper-plasticidad que tiene consecuencias en procesos cognitivos como: percepción, atención, memoria, así como en la emocionalidad y en las funciones mediadas por la neocorteza (Markram & Markram, 2010).

Presumiblemente, las diferencias en la conectividad neuronal modifican las características del flujo de señales a través de estructuras en las redes cerebrales. Estos cambios en el flujo de información podrían resultar en diferencias en la eficacia, organización y contenido de la cognición, las cuales son características de las personas con TEA (Mundy, 2016) que en atención conjunta estarían haciéndose presentes a través de la manifestación atípica de conductas volitivas, la percepción atípica en paralelo y la también atípica resistencia a la distracción (Gernsbacher et al., 2008).

Una situación relativamente típica entre las personas con TEA es que en algún momento han sido pensadas con discapacidad auditiva por su baja respuesta ante estímulos auditivos como su nombre, instrucciones, sonidos, entre otros, lo cual puede ser en realidad uno de los signos de su resistencia a la distracción, es decir, manifiestan una persistencia atípica en la atención por la dificultad de desengancharla de su foco atencional inmediato; también conocida como **hiperfocalización** (Adrien et al., 2001; Garretson et al., 1990; Zwaigenbaum et al., 2005).

Dado que los patrones atípicos de atención y percepción podrían tener implicaciones importantes en el desarrollo y manifestación de las conductas de atención conjunta en personas

dentro del espectro del autismo y sabiendo que dichos patrones pueden cambiar a través del tiempo (Mundy, 2016), abordar su estudio a través de modelos neurocomputacionales que permitan el análisis de las redes de procesamiento a través del tiempo, en las cuales se puedan modelar estilos de percepción y atención atípicas como la hiperfocalización, aportaría conocimiento valioso a la caracterización de la atención conjunta en el TEA.

#### **2.4. Modelos neurocomputacionales y su aplicación a las ciencias cognitivas**

Un modelo es una teoría detallada (Plunkett & Elman, 1997), es la representación de algún hecho o fenómeno que nos permite mostrar las características generales de su estructura, a partir de la cual, nos es posible explicar los mecanismos y procesos implicados y la relación entre ellos, así como develar y comprender los aspectos teóricos que le dan sustento. Todo ello en un proceso de descubrimiento y clarificación.

Los modelos neurocomputacionales están diseñados con el propósito de simular procesos reales en un sistema computacional que nos permita reproducir los componentes importantes del comportamiento o habilidad que queremos estudiar (Munakata & McClelland, 2003; Strube, 2001; Thomas & McClelland, 2008; Tovar, en prensa). Su inspiración se halla en dos nociones importantes: 1) la información y sus propiedades son procesadas de una manera particular por los sistemas neuronales y 2) los principios de dicho procesamiento son susceptibles de ser moldeados a nivel computacional. De tal manera que podríamos considerarlos una aproximación conexionista al entendimiento de la cognición, concibiendo el procesamiento de la información como un conjunto de patrones de activación a través de unidades simples de procesamiento (neuronas) conectadas entre ellas, dando lugar al aprendizaje a partir de la modificación en la fuerza de sus conexiones (Thomas & McClelland, 2001).

Una pregunta que surge naturalmente ante el uso de esta metodología en las ciencias del comportamiento y cognición humana es ¿por qué una computadora podría ser un modelo de ayuda para comprender a un ser humano? Ante lo paradójico que ello parece, es necesario considerar a los modelos computacionales como una metáfora del ser humano, una cuya representación de la realidad dependerá de su apego a la plausibilidad neurológica y psicológica del sistema real cuyo comportamiento pretende simular. Es cierto que no existe garantía de que un modelo computacional pueda ser el espejo del ser humano, sin embargo, al diseñarlo se pretende dotarlo del mismo tipo de información que podría estar disponible para una persona y, por ende, se le da la oportunidad de aprender la tarea cuyo contenido nos interesa analizar (Plunkett & Elman, 1997).

Una metodología basada en este tipo de simulaciones nos permite explorar fenómenos conductuales, evaluar teorías respecto a la ejecución de una tarea experimental, vincular hallazgos de las neurociencias con teorías de la conducta y generar predicciones respecto al comportamiento del modelo ante la modificación de alguna variable o incluso ante afectaciones en él que puedan simular el efecto de alguna lesión o alteración neuronal real (Tovar, en prensa)

Dado que los modelos conexionistas enfatizan el aprendizaje de representaciones internas, es decir, aquellos estados de procesamiento intermedios entre la percepción de la información y la emisión de las respuestas, actualmente es posible estudiar fenómenos del desarrollo e incluso aspectos de los orígenes del conocimiento (Plunkett & Elman, 1997) y de habilidades cognitivas tales como la memoria, la atención, la percepción, la acción, el lenguaje, la formación de conceptos y el razonamiento (Houghton, 2005).

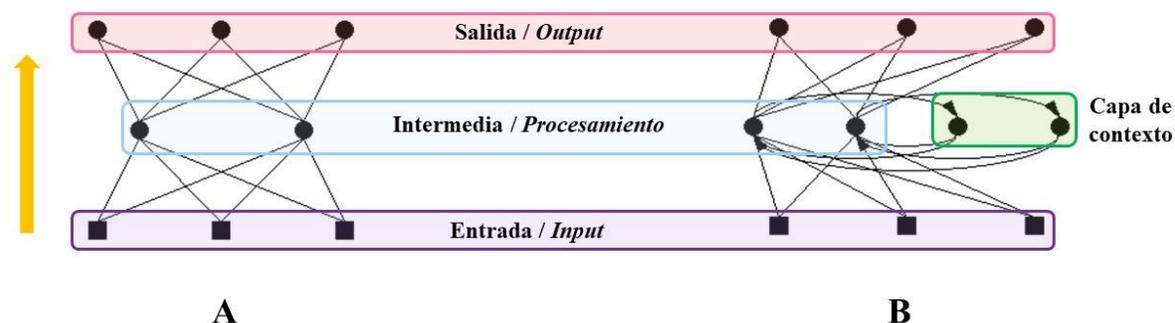
También conocidos como Redes Neuronales Artificiales (ANN, por sus siglas en inglés) o Modelos de Procesamiento Distribuido en Paralelo (PDP), estos sistemas están compuestos de

neuronas artificiales capaces de integrar la información que entra a la red, generar un nivel de activación y también uno de propagación o salida; todo ello gracias a las conexiones que existen entre ellas.

A la conjunción del número de neuronas que conforman una red y la dinámica de sus conexiones la conocemos como **arquitectura**, la cual resulta de gran relevancia en tanto que determina el tipo de información que la red es capaz de procesar y, por ende, el tipo de tarea que puede simularse. Al respecto, existen dos arquitecturas importantes para el presente trabajo y que han sido de amplio uso en el campo del modelamiento computacional de procesos cognitivos: **redes multicapa o tricapa** (o *three-layer feedforward networks* en inglés) y **redes recurrentes** (o *simple recurrent networks* en inglés), cuya revisión resulta imprescindible para comprender la metodología del actual proyecto.

Una red tricapa (Rumelhart et al., 1986) está compuesta por tres capas de unidades: la capa de entrada o *input*, la capa intermedia y la capa de salida o *output*, a través de las cuales la información fluye de manera unidireccional desde la entrada hasta la salida. Por otro lado, una **red recurrente** (Elman, 1990) parte de una arquitectura similar a la de las redes tricapa, pero añade otra capa igual a la intermedia en número de unidades (ver Figura 1). La capa añadida es conocida como **capa de contexto** y tiene una conectividad recurrente con la intermedia que permite hacer reverberar los valores de la activación de la capa intermedia y, con ello, la información previamente introducida a la red se mantiene y es ponderada con la nueva información que entra a la red con el siguiente patrón de *input*. De esa manera, las redes recurrentes incorporan procesos adicionales de memoria inmediata (por mantener en la red una traza de la información presentada en el paso temporal inmediato anterior) y son sensibles a modelar procesos donde el paso del tiempo es relevante (dado que las respuestas o *outputs* de la

red en cada momento son modulados por lo que ha pasado recientemente). En pocas palabras, con las redes recurrentes es posible representar discretamente el tiempo y las secuencias informativas que en él tienen lugar, mediante el uso de los enlaces recurrentes que dotan a la red de una memoria dinámica.



**Figura 1.** Estructura de una red tricapa (A) y una red recurrente (B)

Utilizar a los modelos neurocomputacionales como metodología de investigación en las ciencias cognitivas y del comportamiento nos permite hacer explícitos los procesos relevantes de manera detallada, simular el desempeño de sujetos con lesiones neuronales o funcionamiento neurobiológico atípico, variar parámetros de interés a través de un largo rango de posibilidades, predecir resultados cognitivos y conductuales ante diversas manipulaciones y situaciones, hallar nuevas ideas experimentales para llevar a cabo con humanos y probar teorías difíciles de evaluar en el mundo real. Ante situaciones complicadas como la actual pandemia, por lo tanto, este paradigma resulta de indiscutible utilidad.

### 3. PLANTEAMIENTO DEL PROBLEMA

Desde el final de la década de los 70's, un gran número de investigadores/as han mostrado interés por crear un marco teórico sólido y confiable sobre la atención conjunta y todo lo que ella involucra: las preferencias visuales, coordinación óculo-motora, bases neuroanatómicas, conductas de atención conjunta, modalidades, propiedades de los estímulos en la conducta de IJA, procesos de aprendizaje, entre otros. Esto ha logrado esbozar nociones respecto a la trayectoria de desarrollo de esta habilidad tanto en niños/as con desarrollo típico como en quienes presentan condiciones de desarrollo atípico como TEA.

Con base en la investigación que se ha realizado en el área, han surgido trabajos como el de Bianco et al. (2018) con un enfoque evolutivo y longitudinal, en donde se realiza un esfuerzo por conjuntar los datos generados durante estas cuatro décadas de labor y proponer teorías como la del **Estrechamiento perceptual del seguimiento de la mirada** con una base situada en la postura neuroconstructivista. No obstante, una gran limitante para el establecimiento de una descripción certera del desarrollo y manifestación de la atención conjunta, es la gran variación entre las metodologías que han sido implementadas, considerando la situación de estudio (naturalista/experimental), el paradigma (de seguimiento de mirada “*gaze following*” / de indicación de la mirada “*gaze cueing*”), de los estímulos (sociales: personas reales en interacción, ya sea investigadores/as o cuidadores/as; fotos y/o videos de personas y no sociales: flechas y configuraciones de pixeles) y, de mayor relevancia para este proyecto, las propiedades informativas de los estímulos que señalizan la orientación, siendo las principales variables manipuladas en estudios previos el movimiento y la dirección, sin que sea claro cuál de ellas es atendida preferentemente, o si la atención a estas es dinámica y dependiente del desarrollo.

Además de la imposibilidad para considerar las investigaciones transversales como fotografías confiables de lo que podría ser una trayectoria completa del desarrollo de la atención conjunta por ciertas incompatibilidades en los resultados, el trabajo en seguimiento de mirada (RJA) ha buscado resolver la incógnita del papel de las características dinámicas de los estímulos (dirección y movimiento) en el desencadenamiento de cambios atencionales reflejos mediante la manipulación de estas propiedades en las tareas experimentales, enfrentándose a los retos propios de aislar elementos que en situaciones naturalistas se encuentran siempre juntos, mostrando así un éxito parcial y brindando resultados en el análisis de dirección y movimiento que entre algunos estudios resultan contradictorios.

#### 4. JUSTIFICACIÓN Y OBJETIVOS

La atención conjunta es una habilidad social de irrefutable importancia en el desarrollo de los seres humanos, la cual se manifiesta de manera atípica en personas con Trastorno del Espectro Autista. Dichas atipicidades, en la mayoría de los casos, dificultan el desarrollo de otras habilidades importantes como la comunicación referencial, la teoría de la mente y la adquisición del lenguaje, razón por la que su aprendizaje ha devenido uno de los objetivos principales en el trabajo terapéutico con dichas personas.

Si bien en las últimas décadas se han desarrollado numerosas investigaciones con el objetivo de caracterizar de forma amplia y certera la atención conjunta, el trabajo en población con TEA ha sido menor respecto al realizado con personas con desarrollo típico. Las intervenciones terapéuticas se han basado, primordialmente, en los datos de DT y la posible trayectoria de TEA que ha podido esbozarse, considerando en muy pocas o nulas ocasiones a los mecanismos compensatorios que puedan existir y potenciarse para llegar a resultados clínicos eficientes sin necesariamente buscar igualar la trayectoria normativa.

Una investigación minuciosa de las propiedades informativas y dinámicas de los estímulos que actúan como señales en la IJA y RJA, su percepción por las personas con TEA y su manifestación en las ventanas temporales del desarrollo en TEA resulta de vital importancia para fines básicos y aplicados en el campo del autismo.

Adicionalmente, las limitaciones que impuso el confinamiento por la pandemia de COVID-19, con la cual no fue posible por razones éticas, prácticas y de salud el realizar estudios con infantes con desarrollo típico y con TEA, motivaron un acercamiento computacional que nos permitiera simular los componentes bajo análisis de la atención conjunta, crear patrones de estimulación que emulen la información cuya disponibilidad pensamos que es plausible para los

humanos, crear una correspondencia cercana entre los comportamientos y restricciones de procesamiento incorporadas en el modelo, para comprobar o refutar las hipótesis establecidas y generar predicciones que puedan después ponerse a prueba en investigaciones empíricas.

### **Objetivo general**

- Crear un modelo computacional capaz de simular el desempeño de participantes en tareas de seguimiento de mirada frente a diferentes tipos de estímulos (sociales: mirada y cabeza; no sociales: líneas) con diferentes propiedades dinámicas (movimiento y dirección final).

### **Objetivos específicos**

- Explorar el valor informativo de pistas sociales y no sociales para analizar su efecto en el desarrollo de la RJA.
- Proponer una trayectoria de la RJA en trayectorias típicas del desarrollo y en el Trastorno del Espectro del Autismo.
- Simular diferencias de procesamiento de la información y evaluar su impacto en el desarrollo de la RJA.
- Encontrar evidencia relacionada con los hallazgos y propuestas existentes respecto al surgimiento y desarrollo de la RJA.
- Investigar qué tipo de pistas y propiedades informativas resultan más relevantes y en qué momento para el desarrollo de la RJA en TEA y DT.

## 5. HIPÓTESIS

- H1.** Los rostros, ojos y boca son señales sociales que varían en la cantidad de información que brindan (de más a menos, respectivamente) y las señales visuales más informativas permiten respuestas más tempranas y eficaces a la atención conjunta.
- H2.** El movimiento percibido de estímulos visuales es el primer facilitador para el desarrollo de la RJA en población con desarrollo típico.
- H3.** En la población con desarrollo típico la RJA se manifiesta con mayor facilidad (menor latencia y mayor acierto) ante estímulos sociales.
- H4.** El desarrollo de la RJA en TEA presentará una trayectoria distinta a la del desarrollo típico, pudiendo estar solamente demorada, atípica o demorada y atípica.

## 6. MÉTODO

El acercamiento metodológico estuvo basado en desarrollar modelos conexionistas (Thomas & McClelland, 2001) de desarrollo típico y de autismo para la simulación de tareas de seguimiento de mirada como paradigmas de Respuesta a la Atención Conjunta.

Para cada población simulada se crearon 2 arquitecturas capaces de simular el procesamiento de la dirección (arquitectura 1) y del movimiento (arquitectura 2) de estímulos ambientales como pistas iniciadoras de la atención conjunta.

Para la presentación de estímulos se consideraron 3 condiciones distintas que fueron la simulación de 1) la cabeza, 2) la mirada y 3) una pista no social. Se realizaron simulaciones de desarrollo típico (redes 1-3) y simulaciones con desarrollo atípico que aproxima TEA. Como se describirá más adelante, se exploró la simulación de TEA con 3 manipulaciones distintas, lo que resultó en un total de 4 tipos de población simulada: una con desarrollo típico y tres con TEA.

La combinatoria entre poblaciones simuladas (4), arquitecturas (2), y tipo de estímulos (3) generó un total de 24 condiciones o redes distintas.

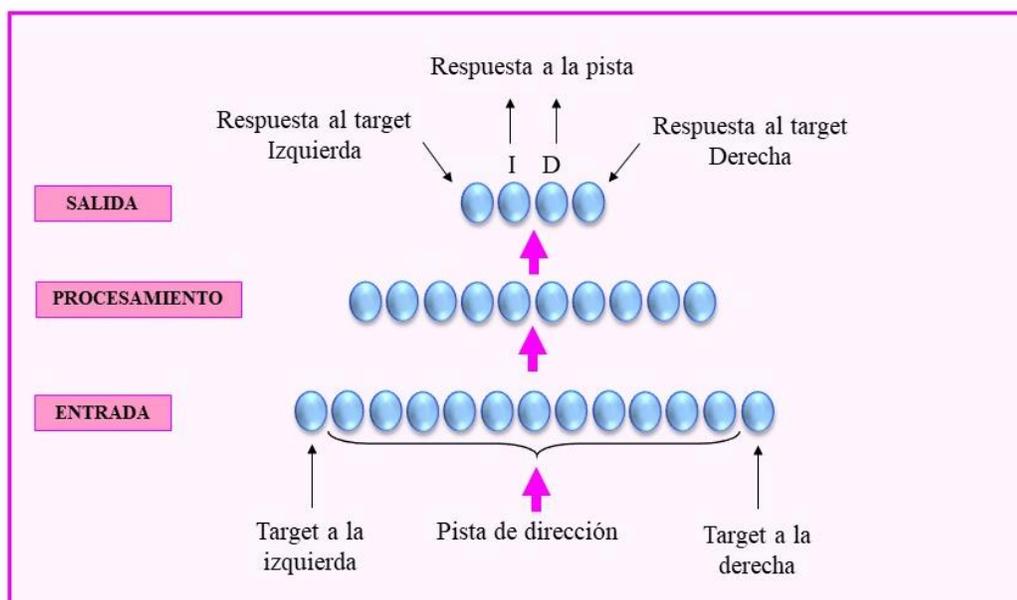
Para facilitar la descripción de las simulaciones implementadas, primero describiremos las diferencias entre las dos arquitecturas que permitieron generar simulaciones de dirección (DIR) y de movimiento (MOV). Posteriormente, se detallará la diferencia entre el tipo de estímulos y, finalmente, explicaremos cómo se desarrollaron las simulaciones de la población con TEA.

### 6.1. Redes Neuronales para Simulaciones de Dirección (DIR)

La arquitectura básica sobre la cual se desarrollaron todas las simulaciones es una tradicional en la perspectiva conexionista (Plunkett & Elman, 1997; Thomas & McClelland, 2001) compuesta por tres capas: una primera capa de entrada (*input layer* en inglés) donde se

activan directamente neuronas para representar distintos arreglos de estimulación; una segunda capa (*hidden layer* en inglés) encargada de realizar representaciones intermedias entre los estímulos y las respuestas; y finalmente, una capa de salida o respuesta (*output layer* en inglés) en donde la activación de las neuronas se interpreta como las respuestas del modelo.

Para las simulaciones de Dirección se realizaron “corridas” del modelo donde la información fluyó desde la capa de entrada hasta la de salida (ver Figura 2). Posteriormente se detallará cómo fue la presentación de estímulos en la capa de entrada y el análisis de los patrones de activación de la capa de salida.

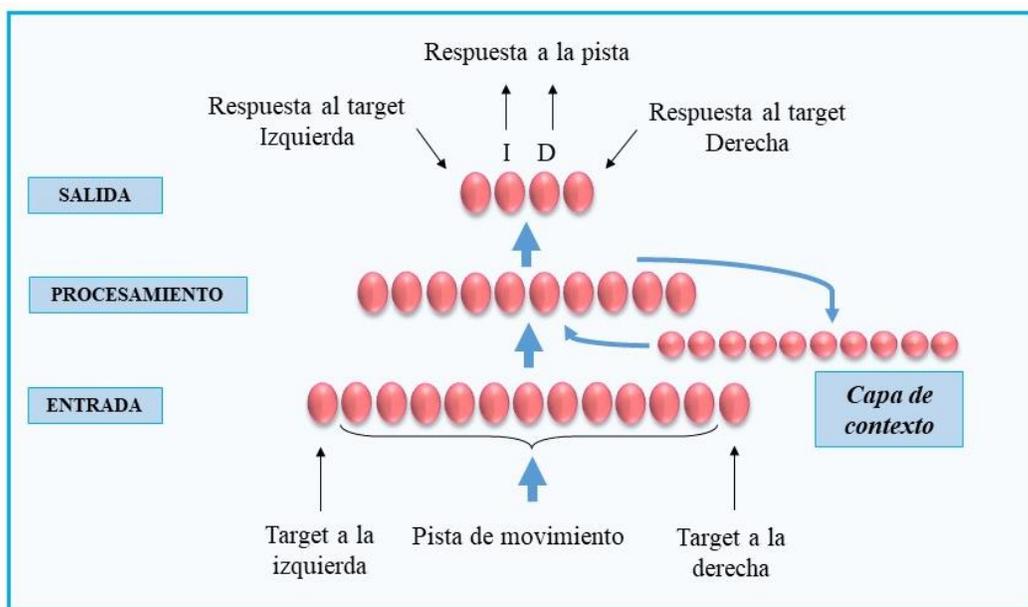


**Figura 2.** Red neuronal para simulación de Dirección

## 6.2. Redes Neuronales para Simulaciones de Movimiento (MOV)

Dado que el movimiento es definido como un cambio de la posición de un cuerpo a lo largo del tiempo respecto a un sistema de referencia, el modelamiento requirió de una arquitectura capaz de realizar un procesamiento de dicha información dinámica, y de guardar

memoria secuencial de lo ocurrido antes de cada nueva entrada de información, para lo cual se usó una Red Neuronal Recurrente (o RNN por sus siglas en inglés) (Elman, 1990). Esta arquitectura es muy similar a la tradicional de tres capas, con la particularidad de contar adicionalmente con una capa de contexto (*context layer*) compuesta por unidades que funcionan como una memoria interna de la red al evaluar y ponderar las propiedades de cada entrada de información con respecto a los datos del paso inmediato anterior (ver Figura 3).



**Figura 3.** Red neuronal para simulación de Movimiento

Para las simulaciones de Movimiento, por lo tanto, la información fluyó desde la entrada hacia la capa de procesamiento intermedia y de esta hacia la capa de salida, por una parte, y hacia la capa de contexto por otra. Finalmente, la capa de contexto tiene conexiones recurrentes con la capa intermedia como se aprecia en la Figura 3, lo que permite ponderar la información recibida anteriormente.

Cada una de las arquitecturas estuvo compuesta por 13 unidades de entrada en donde podía simularse la aparición tanto de las pistas visuales como del objeto target, que son los

estímulos relevantes en las tareas de seguimiento de mirada; otras 10 unidades de procesamiento intermedias (y 10 unidades de contexto adicionales para la arquitectura recurrente que permite simular movimiento); y 4 unidades de salida para representar las respuestas visuales sobre las secciones izquierda y derecha del ambiente. Se designaron 4 unidades de salida para que las dos más externas (izquierda y derecha) representaran las respuestas al *target* que aparecería a la izquierda o derecha en cada ensayo, mientras que las dos unidades internas (izquierda y derecha) representaran las respuestas de seguimiento de la pista que aparece en cada ensayo, como se detalla más adelante. Esto permite capturar el componente de respuesta durante las tareas de seguimiento de mirada, que consiste en observar y atender a la pista que aparece centrada, para luego observar y atender el objeto *target* que aparece en algún extremo.

### **6.3. Patrones de Estimulación**

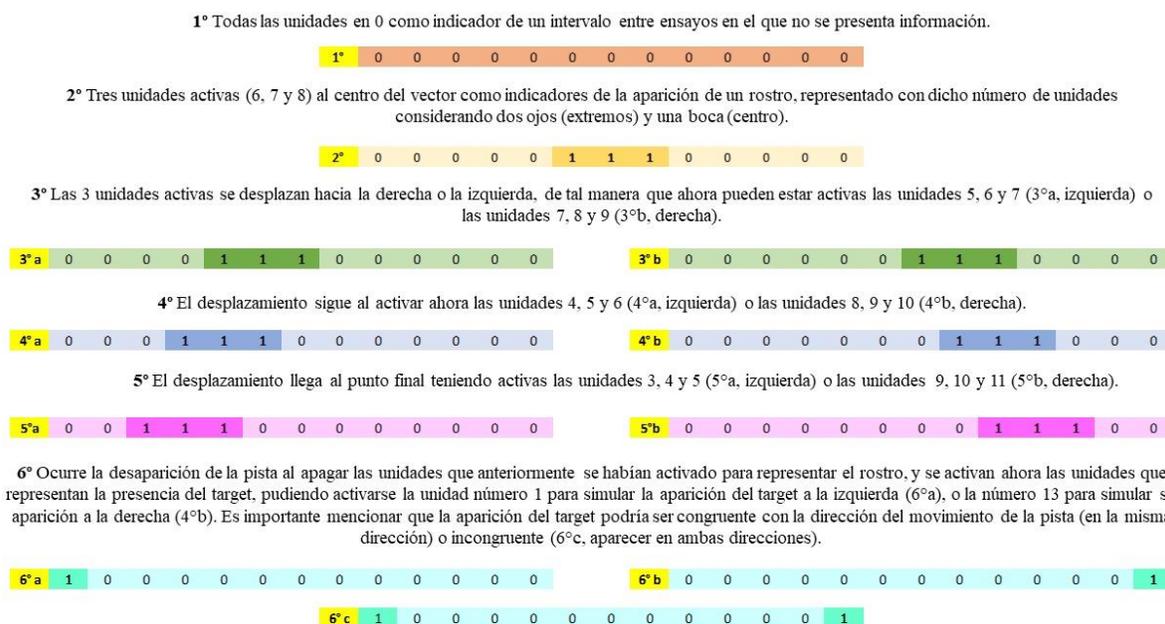
A través de un análisis de la tarea de seguimiento de mirada, diseñada para computadora por Hood y colaboradores (1998), se identificaron los elementos fundamentales de la misma con la intención de crear los patrones de estimulación, de tal manera que estos debían poseer, tal como en la tarea original: una pista que aparece al centro del espacio y desde ahí se mueve o se dirige hacia la izquierda o derecha anticipando la aparición del objeto *target* que se posiciona en el área izquierda o derecha del espacio. Con el propósito de facilitar el desenganche de la atención de cada estímulo, la pista y el *target* son siempre elementos aislados entre sí que no aparecen al mismo tiempo en ningún momento.

La estimulación de las redes se llevó a cabo mediante bloques de vectores cuyos valores se diseñaron para capturar diferencialmente la cantidad de información implicada en cada uno de los tres tipos de condiciones estimulares simuladas (i.e., cabeza, mirada y no-social).

Adicionalmente, las simulaciones de movimiento (MOV) requirieron de más vectores por ensayo simulado para capturar el carácter dinámico del movimiento.

Para la simulación de las condiciones de *cabeza*, *mirada* y *no-social*, se asumió que la diferencia más relevante entre ellas residía en la cantidad de información que cada tipo de estímulo provee. Con base en este supuesto, y sabiendo que los resultados y conclusiones quedarán restringidas/anclada(s) por esta suposición, consideramos para esta primera exploración que la cabeza sería representada por 3 bits, la mirada por 2 bits y lo no social por 1 bit.

Con el fin de explicar la composición de los patrones de estimulación, en la Figura 4 se describe uno de los patrones de estimulación más complejos, que fue el de la condición de cabeza en la situación de movimiento (i.e., MOV\_cabeza). Puede apreciarse que este ensayo estuvo compuesto por un bloque de 6 vectores que aparecen aislados en la figura con el fin de acompañarlos con su descripción.



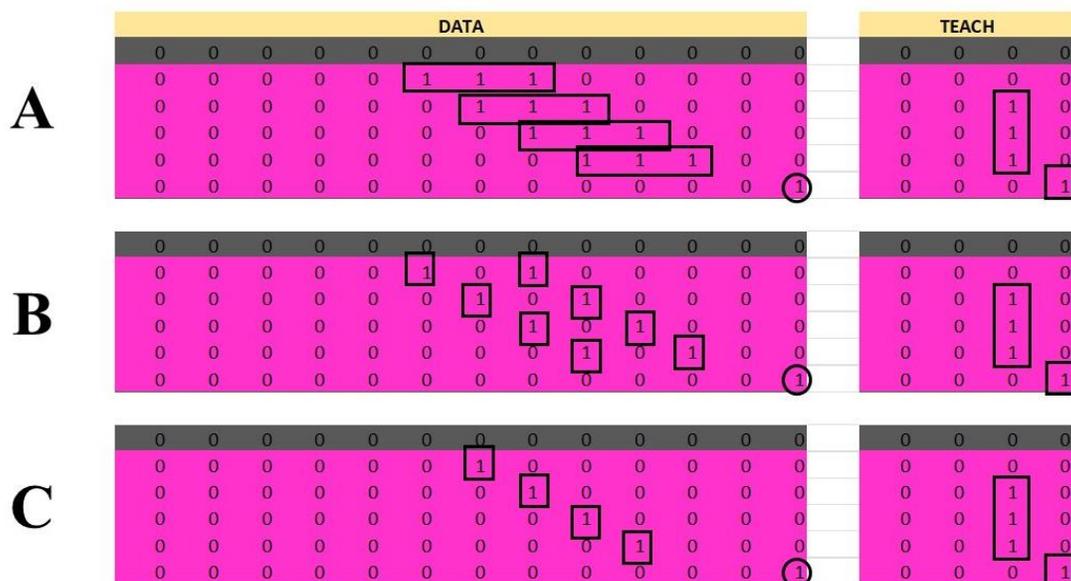
**Figura 4.** Descripción gráfica y textual de un bloque de vectores para simular ensayos de movimiento de cabeza

Ante el patrón de estimulación mostrado en la Figura 4, la red fue entrenada para brindar una respuesta representada en 4 unidades, de tal manera que la activación en las unidades 1 y 4 simularían la atención hacia la aparición del target a la izquierda (unidad 1) o a la derecha (unidad 4), mientras que las unidades 2 y 3 se usaron para simular las respuestas de reconocimiento de la dirección de la pista, pudiendo encontrarse o dirigirse hacia la izquierda (unidad 2) o hacia la derecha (unidad 3).

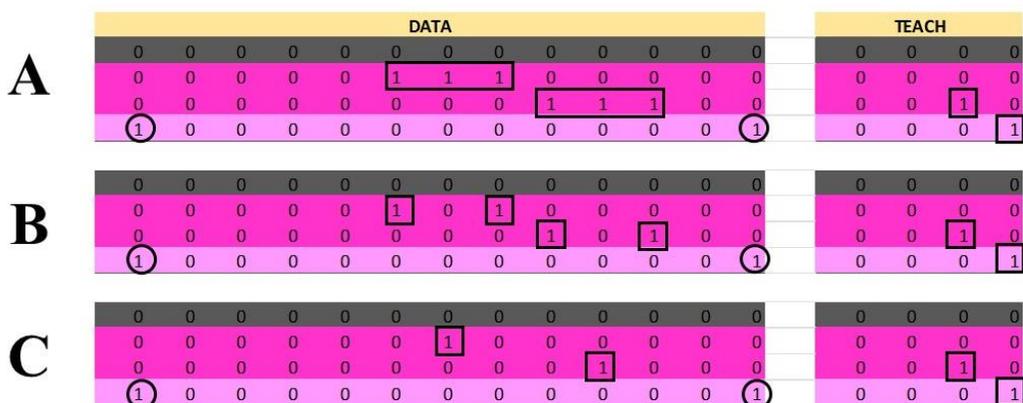
Los patrones de activación/estimulación para las demás condiciones del estudio fueron creados bajo los mismos principios y estructura, siendo el número de unidades activas en la capa de entrada para la pista, la diferencia principal entre las condiciones de *cabeza*, *mirada* y *no-social*. Por ejemplo, en MOV\_mirada, el 2º vector presentó activación solo en dos de las unidades (6 y 8, simbólicamente una por cada ojo), desplazándose de la misma manera hacia la derecha o hacia la izquierda en los 3 renglones siguientes. En MOV\_no-social, la pista estuvo representada solo por una unidad, de tal manera que el 2º vector tuvo la unidad 7 activa y esta se fue desplazando hacia la derecha o izquierda en los 3 vectores siguientes. El contraste entre estas 3 condiciones, capturado en los vectores de estimulación, puede apreciarse en la Figura 5.

Para las simulaciones de dirección, donde no se incluyó movimiento de las pistas, el diseño de vectores fue realizado de manera muy similar, aunque el movimiento fue eliminado al presentar a las redes únicamente los vectores 1, 2, 5 y 6, brindándoles solamente información del intervalo entre ensayos, la posición inicial de la pista, la posición final de la pista y el lugar de aparición del *target* (ver Figura 6).

Las simulaciones se realizaron siguiendo las fases tradicionales de entrenamiento y prueba que se realizan en estudios empíricos como se describe a continuación.



**Figura 5.** Patrones de estimulación para las condiciones MOV\_cabeza (A), MOV\_mirada (B) y MOV\_no-social (C) en bloques congruentes hacia la derecha, donde DATA representa la información de entrada (*input*) y TEACH la respuesta (*output*) entrenada.



**Figura 6.** Patrones de estimulación para las condiciones DIR\_cabeza (A), DIR\_mirada (B) y DIR\_no-social (C) en bloques incongruentes con aparición de un target en cada lado, donde DATA representa la información de entrada (*input*) y TEACH la respuesta (*output*) entrenada.

#### 6.4. Entrenamiento

Cada una de las condiciones bajo estudio fue entrenada durante 5000 ensayos, cantidad acordada con base en el comportamiento de las redes y el alcance de su mayor aprendizaje hasta no mostrar variación significativa. Es importante recordar que cada ensayo requirió de 6 vectores de estimulación en el caso de movimiento, o de 4 vectores de estimulación en el caso de dirección. El entrenamiento se programó determinando el número necesario de *sweeps* o corridas para cada arquitectura. Un *sweep* corresponde con la presentación de un solo vector de estimulación a la red neuronal. De esta manera, las condiciones de movimiento fueron entrenadas por 30000 corridas o *sweeps* (i.e.,  $30000/6 = 5000$  ensayos) y las condiciones de dirección fueron entrenadas por 20000 corridas o *sweeps* (i.e.  $20000/4 = 5000$  ensayos). Debe considerarse que en estas simulaciones una corrida no es equivalente a un ensayo con un/a participante humano/a, sino a una iteración de la información que la red utiliza, de tal forma que un ensayo con un/a participante humano/a suele capturarse con numerosas corridas o *sweeps* de una red neuronal.

La matriz creada para cada condición contó con un 50% de ensayos hacia la izquierda y un 50% hacia la derecha. Del total de ensayos hacia cada dirección, el 20% fueron incongruentes, es decir, el *target* aparecía en ambas direcciones y no solo en la posición anticipada por la pista.

En todas las condiciones descritas, el entrenamiento se realizó de la siguiente manera: los patrones de estimulación que representaron cada uno de los ensayos fueron presentados en la capa de entrada. Las activaciones fluyeron por la red hasta alcanzar la capa de salida y los valores alcanzados en esta fueron evaluados como la respuesta de la red. Cada valor de activación en las 4 unidades de salida de las redes fue contrastado contra el valor “esperado” u

objetivo. La diferencia entre el valor de activación y el valor esperado fue en consecuencia el nivel de error. Dicho error es local cuando se calcula por unidad, y es global cuando se calcula como el promedio de las 4 unidades de respuesta en un momento dado. Durante el entrenamiento estuvo activo el algoritmo de aprendizaje por retropropagación o *backpropagation* en inglés (Plunkett & Elman, 1997) que realiza ajustes en todas las conexiones del modelo, comenzando por la última y yendo “hacia atrás” hasta llegar a la primera capa de conexiones, y dichos ajustes están en función de disminuir el valor de error de la capa de salida. En este sentido, los valores de error se interpretan consecuentemente como un valor inverso del nivel de aprendizaje; es decir, a mayor disminución del error, mayor nivel de aprendizaje.

El análisis del entrenamiento se centró en apreciar el comportamiento de la curva de error (o aprendizaje) y realizar las comparaciones y análisis intra e inter-condiciones, considerando el tiempo (cantidad de ensayos o *sweeps*) en el que cada una de las redes aprendió a manifestar una respuesta correspondiente a la RJA; es decir, el momento en que disminuyó su valor de error en la ejecución de la tarea al manifestar la respuesta o *output* esperado.

### **6.5. Modelamiento del Trastorno del Espectro Autista**

Con el objetivo de representar y simular las limitaciones y atipicidades que tradicionalmente se describen en la trayectoria del desarrollo de personas con TEA con impacto en la atención visual conjunta, fueron implementadas restricciones de procesamiento para capturar 2 grandes condiciones: hiperfocalización (HF, Adrien et al., 2001; Garretson et al., 1990; Zwaigenbaum et al., 2005), percepción atípica (PA, Gernsbacher et al., 2008; Markram & Markram, 2010; P. C. Mundy, 2016); así como la combinación de ambas (HF+PA). Es importante notar que las simulaciones computacionales permiten entonces hipotetizar sobre la

posible contribución de cada una de estas condiciones, así como el efecto sumado, sobre las ejecuciones en tareas de seguimiento de mirada.

La hiperfocalización se define como una persistencia atípica en la atención por la dificultad de desengancharla de su foco atencional inmediato (Adrien et al., 2001; Garretson et al., 1990; Zwaigenbaum et al., 2005). Para capturar un tipo de procesamiento hiperfocalizado, forzamos que, en cada momento representado en la simulación, la red se hiperfocalizaría sobre uno de los elementos presentados en la entrada de información (capa de *input*), elemento que era susceptible de cambiar en el siguiente momento. Esta implementación computacional se realizó tomando los vectores de estimulación típicos y seleccionando aleatoriamente una de las unidades activas en cada momento. La unidad seleccionada se mantuvo activa, mientras que el resto de las unidades se apagaron.

Por otra parte, la percepción atípica (PA) en autismo se ha descrito a partir de las diferencias en la conectividad neuronal que modifican las características del flujo de señales a través de estructuras en las redes cerebrales y pueden resultar en diferencias en la eficacia, organización y contenido de la cognición (Gernsbacher et al., 2008; Mundy, 2016). Para modelar esta condición, se introdujo ruido en los vectores de estimulación de la capa de entrada, de manera que las redes fueron forzadas a procesar información que se recibió incompleta e inconsistente a través de los ensayos del entrenamiento. El ruido fue introducido al tomar los patrones típicos de activación y restar un valor decimal aleatorio a cada una de las unidades activas. El ruido no fue constante, sino que, para cada momento simulado en los patrones de activación, se introdujo un nivel distinto (aleatorio) de ruido en las unidades activas. Finalmente, para contar con simulaciones que implementaran las dos atipicidades en el procesamiento, características del TEA, las redes de la condición HF+AP se vieron sometidas a ambas

manipulaciones, activando una sola unidad de la capa de entrada y añadiendo valores de ruido a esta en cada ensayo (Ver Figura 7).

	DATA												TEACH					
<b>A</b>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
<b>B</b>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.183	0.737	0.884	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.133	0.571	0.787	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.630	0.580	0.334	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.325	0.586	0.154	0.000	0.000	0.000	0.000	0.000	0.904	
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
<b>C</b>	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.920	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.510	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.780	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.790	

**Figura 7.** Patrones de estimulación para condición de movimiento de cabeza congruente hacia la derecha en las manipulaciones de TEA.

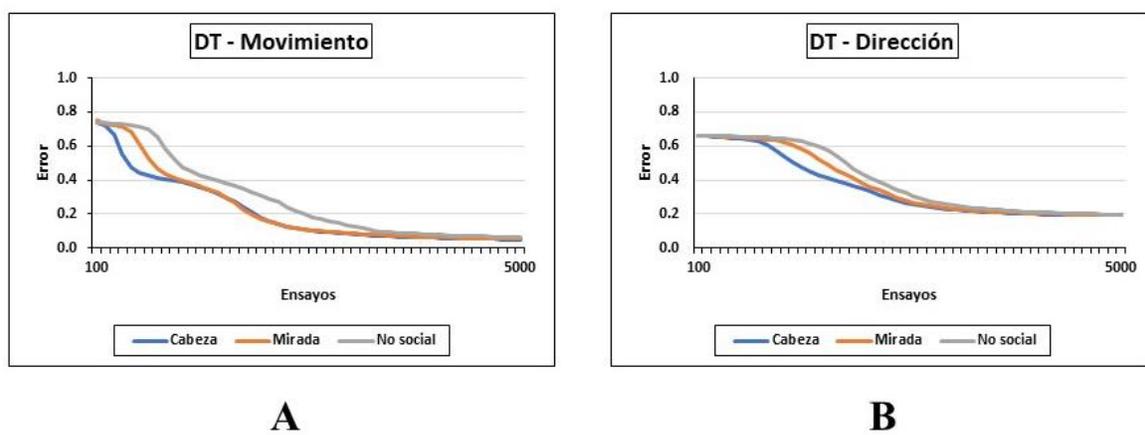
**A.** TEA\_MOV\_cabeza\_HF, **B.** TEA\_MOV\_cabeza\_PA y **C.** TEA\_MOV\_cabeza\_HF+AP, donde DATA representa la información de entrada (*input*) y TEACH la respuesta (*output*) entrenada.

## 7. RESULTADOS Y DISCUSIÓN

El análisis de los resultados se basó en las gráficas de error generadas por las poblaciones y condiciones simuladas. El objetivo fue describir el desempeño de las redes en la tarea a partir de la disminución del error, así como tener un acercamiento con una perspectiva de desarrollo mediante la identificación de los momentos en los que se generan cambios importantes en el nivel de error o aprendizaje.

### 7.1. Desarrollo típico (DT)

Para comenzar, abordaremos los resultados respecto a cada tipo de estímulo en las simulaciones de desarrollo típico: **DT\_MOV** y **DT\_DIR** (ver Figura 8).



**Figura 8.** Gráficas de error en simulaciones de **Desarrollo Típico** en condición de Movimiento (A) y Dirección (B).

En la condición de movimiento **DT\_MOV**, se observa una primera e importante disminución del valor de error en la simulación de **Cabeza** (línea azul) que comienza temprano en el ensayo 300 con un error de **0.73** hasta alcanzar un valor de **0.42** al llegar al ensayo 700, mostrando un primer descenso de la curva. A esta pendiente le sigue otra que comienza en el

ensayo 1100 con un error de **0.38** y llega al **0.14** en el ensayo 2200, de tal forma que ambas disminuciones resultan en un comportamiento de doble descenso o descenso bifásico.

La misma trayectoria bifásica es observada en la simulación de *Mirada* (línea anaranjada) pero notablemente se aprecia con una demora en comparación con la simulación de cabeza; la primera caída del error comienza hasta el ensayo 500 con un valor de **0.72** y termina en el ensayo 900 llegando a un error de **0.43**. La segunda caída comienza en el ensayo 1200 **0.37** y alcanza un error de **0.12** al llegar a los 2400 ensayos.

Las trayectorias de disminución del error en las simulaciones de desarrollo típico son parecidas para las condiciones en las que las pistas son Cabeza y la Mirada, siendo la principal diferencia entre estas, que el nivel de disminución del error y, en consecuencia, el aprendizaje, son más tempranos en la condición de Cabeza; esto se explica en nuestras simulaciones a partir de una mayor cantidad de información en esta condición comparada con la condición de Mirada.

Notablemente, la tendencia de disminución del error es diferente en la simulación de un estímulo *No social* (Fig. 8, panel A, línea gris), observando que la trayectoria de caída en el valor de error es continua y no bifásica, encontrándose en un **0.69** cuando la red ha sido entrenada por 700 ensayos y alcanzando un valor de **0.41** en el ensayo 1400, de donde posteriormente el descenso es continuo, distinto al comportamiento bifásico de las simulaciones de Cabeza y Mirada.

La presencia de fases en el desarrollo de la Atención Conjunta fue planteado hace ya 30 años por Butterworth y Jarret (1991), al proponer 3 mecanismos de procesamiento espacial, cuya evolución se encuentra estrechamente ligada con las características manifiestas de la Atención Conjunta. Inicialmente, el mecanismo *ecológico* permite a la pista indicar la dirección general hacia la cual mirar; posteriormente, el desarrollo cognitivo permite al mecanismo *geométrico*

surgir y brindar la capacidad de extrapolar una trayectoria relativamente correcta desde la pista de una persona adulta hacia la ubicación específica del target en el espacio visual; finalmente, el mecanismo *representacional* establece una extensión de la referencia conjunta hacia lugares que se encuentran fuera del campo visual.

El comportamiento bifásico de la disminución del error en la simulación de los estímulos sociales (Cabeza y Mirada) en la condición de movimiento de nuestro modelo podría estar relacionado con la propuesta de Butterworth y Jarret, no obstante, nuestra manipulación de la dimensión, información dinámica y el supuesto de fondo del valor social de la pista, expande el desarrollo por fases más allá del procesamiento espacial. De esta manera, el desarrollo de la RJA podría también depender en gran medida del valor social de la pista, capturado aquí en términos de un componente “más informativo”. Esto nos permite invitar a una mayor investigación sobre la relación entre los efectos que tienen pistas sociales informativas y el consecuente desarrollo de la teoría de la mente, es decir, la capacidad de atribuir pensamientos e intenciones a las personas que nos rodean (Baron-Cohen, 1995a), dando cuenta de procesos de desarrollo psicosocial como es, por ejemplo, descubrir que cuando una persona manifiesta un cambio en la dirección de su mirada (o en su postura), tiene la intención de compartir con otra persona su foco atencional.

Tomando en consideración el papel de la teoría de la mente (Baron-Cohen, 1995a) y de los mecanismos de procesamiento visoespacial (Butterworth & Jarrett, 1991), el primer descenso del error en la red de movimiento en desarrollo típico podría capturar aspectos de la orientación refleja hacia un cambio motor, dinámico y grueso (Anstis et al., 1969) en el ambiente, lo cual permite que el mecanismo ecológico surja y de paso a los siguientes; mientras que el segundo podría capturar aspectos de descenso que representaría el aprendizaje y manifestación de una RJA después de haber experimentado el valor social de orientarse hacia donde otra persona lo

hace y saber que algo puede ocurrir en ese lugar, como lo refuerza el aprendizaje supervisado de nuestra red ante respuestas correctas.

Ante este análisis de nuestro modelo no solo apoyamos el papel del procesamiento visoespacial y su desarrollo por fases en la manifestación de la RJA, sino también consideramos viable la idea del aprendizaje de la RJA como producto de un proceso de reforzamiento (Corkum & Moore, 1995, 1998; Moore et al., 1997) que implica factores ambientales (la cantidad de información) y precursores ontogenéticos relacionados particularmente con el procesamiento de la información ambiental disponible (teoría de la mente y mecanismos de procesamiento visoespacial).

En la condición de dirección **DT\_DIR**, la trayectoria de aprendizaje ante los tres tipos de estímulo fue muy similar a lo descrito para movimiento, difiriendo principalmente en el tiempo o la cantidad de ensayos sucedidos antes de comenzar los descensos importantes en el valor de error (Fig. 8, panel B). Las tres condiciones comenzaron con un valor de **0.66**, siendo de nuevo la condición de cabeza en la que comenzó a disminuir el error primero en el ensayo 900; para la condición de mirada, esto sucedió en el ensayo 1200 y en la condición de estímulo no social hasta los 1500 ensayos. Mientras los estímulos sociales alcanzaron un error de **0.21** a los 3500 ensayos, la simulación de estímulo no social alcanzó el mismo valor a los 3700 ensayos.

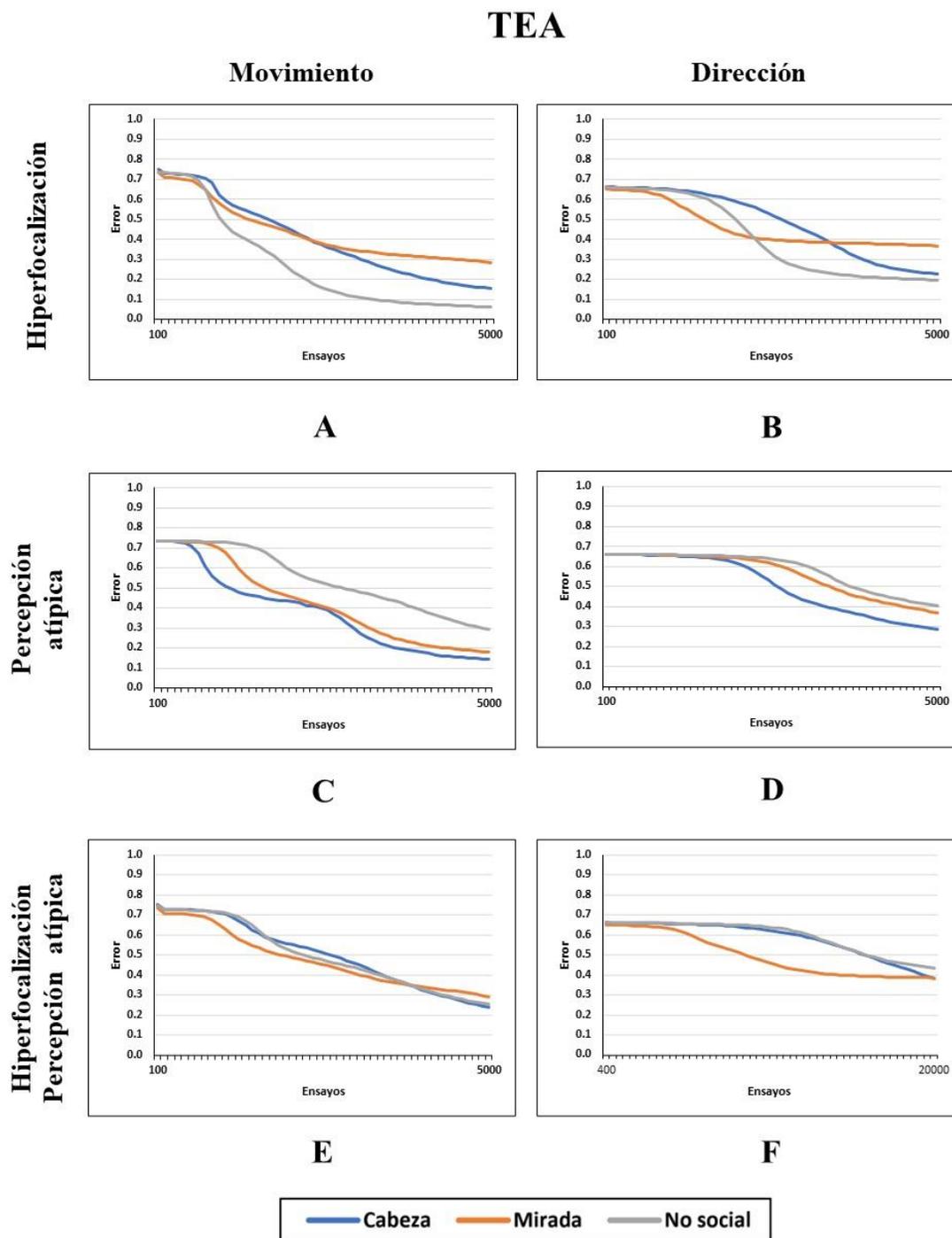
Ante la ausencia de movimiento, ninguna condición en **DT\_DIR** tuvo un comportamiento bifásico como el observado en **DT\_MOV**, lo que notablemente apoya la hipótesis del movimiento como el primer gran facilitador para el desarrollo de la Respuesta a la Atención Conjunta (Corkum & Moore, 1995; Farroni et al., 2000). La comparación entre condiciones dinámicas y tipo de estímulo bajo los supuestos de cantidad de información como variables independientes nos permite sostener la hipótesis de que el movimiento percibido (inicialmente de

Cabeza y luego de Mirada) apoya la emergencia de una Respuesta a la Atención Conjunta durante los primeros meses de vida. No obstante, la transición entre la percepción inicial de la Cabeza y el Movimiento y posteriormente de la Mirada y la Dirección no es considerada como un remplazo de la dirección al movimiento y de la mirada a la cabeza como pistas relevantes para el desarrollo de la RJA, sino como la capacidad de procesar pistas más finas y menos informativas gracias a la experiencia perceptual previa con pistas más gruesas y más informativas, un apoyo al planteamiento del estrechamiento perceptual en el seguimiento de mirada (Del Bianco et al., 2018).

## **7.2. Trastorno del Espectro del Autismo (TEA)**

### **7.2.1. Hiperfocalización (HF)**

Notablemente, en la condición **TEA\_MOV\_HF** aparece por primera vez en estas simulaciones una ventaja del estímulo no social (Fig. 9, panel A, línea gris) en el descenso del error en la RJA; el nivel inicial de error para esta pista comienza en **0.73** y disminuye mínimamente después de 800 ensayos para llegar a **0.64**, momento en el que la red manifiesta un mayor aprendizaje a través de una caída pronunciada del error para llegar a un valor de **0.43** solo 400 ensayos después (i.e. en el ensayo 1200). El aprendizaje comienza a ser gradual y finalmente el error alcanza un valor de **0.06** en el último ensayo del entrenamiento. Por otro lado, la curva de error para esta condición en la pista que simula la mirada (Fig. 9, panel A, línea anaranjada) presenta un comportamiento de disminución continua sin caídas remarcables, comenzado con un valor de **0.73** al inicio del entrenamiento para disminuir solo al **0.2** después de los 5000 ensayos de entrenamiento. Finalmente, la curva que representa el movimiento de la cabeza (Fig. 9, panel A, línea azul) parece una combinación de las dos pistas descritas, en tanto que la mayor parte de



**Figura 9.** Gráficas de error en simulaciones de **Trastorno del Espectro Autista**.

su descenso es lento y continuo, pero llega a presentar una pequeña caída remarcable. De esta manera, el valor de error comienza en **0.75** y desciende lentamente hasta alcanzar **0.68** en el ensayo 900, en donde se observa un declive pronunciado para alcanzar un valor de **0.57** solo 200 ensayos después (i.e. ensayo 1100) y luego tener un descenso gradual hasta alcanzar un valor de **0.15** tras los 5000 ensayos, el cual se termina posicionando en medio de los otros dos tipos de estímulos.

Por otra parte, la condición **TEA\_DIR\_HF** comenzó con un comportamiento diferente a su red par de movimiento y, al llegar al ensayo 3100 comienza a mostrarse de manera similar a ella. El primer descenso en el valor de error se observa para la pista que simula la Mirada (Fig. 9, panel B, línea anaranjada), para la que el error comienza con un valor de **0.64** y se mantiene bastante estable hasta comenzar un descenso en el ensayo 900 para llegar de **0.61** hasta **0.41** en el ensayo 2100, punto en el que ya el aprendizaje de la tarea parece estabilizarse de nuevo terminando el error en **0.36** después de los 5000 ensayos. Con base en el tiempo que toma mostrar un aprendizaje de la tarea (i.e. disminución pronunciada del valor de error), la siguiente pista ante la cual se manifiesta la RJA es ante la no social (Fig. 9, panel B, línea gris); comenzando con un valor de **0.66** y llegando solo a **0.61** después de 1500 ensayos, la curva alcanza un valor de **0.33** en un intervalo de 1000 ensayos (i.e. del ensayo 1500 al 2500) para continuar su disminución de una manera gradual y alcanzar un error de **0.19** al terminar el entrenamiento. Finalmente, la pista que simula la Cabeza se presenta sin variación remarcable (Fig. 9, panel B, línea azul), pero con un descenso tan constante que, incluso comenzando con un valor de **0.66** y mostrando un descenso pronunciado hasta los 2100 ensayos con un error de **0.58**, al final del entrenamiento durante los 5000 ensayos alcanza un valor muy cercano al del estímulo

no social, terminando con un error de **0.22** y posicionándose debajo de la pista que simula la mirada.

Analizando ambas condiciones (movimiento y dirección) para la simulación de **Hiperfocalización** como estilo característico de procesamiento perceptual en TEA, se puede observar que una cantidad de información disminuida debido a la hiperfocalización parece favorecer el inicio y desarrollo tempranos de la RJA. Ante el supuesto de los estímulos no sociales como pistas menos complejas y también menos informativas, consideraríamos que la preferencia por este tipo de estímulos incluso desde la infancia (Klin et al., 2009) en personas con TEA caracterizadas por un estilo de procesamiento con hiperfocalización, podría entenderse como un mecanismo adaptativo mediante el cual la hiperfocalización es compatible con la escasa información proveída por una pista no social, el cual podría evolucionar gradualmente hacia una RJA asociada al movimiento percibido de la cabeza y posteriormente, tomando más tiempo del habitual e incluso pudiendo no alcanzar un aprendizaje del 100%, atender la información dada solamente por la mirada de otra persona.

Ante el aprendizaje de la RJA con estímulos no sociales y la ayuda inicial de un movimiento grueso, vale la pena explorar la función que el movimiento de otras partes del cuerpo podría tener en el alcance de episodios de atención conjunta en infantes con TEA, pues una reciente línea de investigación (Fausey et al., 2016; Yu & Smith, 2013, 2017a, 2017b) postula la posibilidad de establecer atención conjunta incluso sin seguimiento de mirada, coordinando la atención visual hacia los objetos a través de la coordinación ojo-mano . En este caso, sería valioso reconceptualizar la RJA e investigarla más allá de un paradigma de seguimiento *de* la mirada para transitar a un seguimiento *con* la mirada (incluso con otros

sentidos) o un *seguimiento atencional* de aspectos más globales como la postura corporal y el movimiento y dirección de las manos.

### 7.2.2. *Percepción atípica (PA)*

Las curvas de error de las condiciones **TEA\_MOV\_PA** y **TEA\_DIR\_PA** tienen un comportamiento idéntico al de las redes de Desarrollo Típico para movimiento y dirección, respectivamente, manifestando como única particularidad una demora en su evolución.

De esta manera, para la red de **movimiento** ante **percepción atípica**, el valor de error para los tres tipos de estímulo comenzó en **0.73** y mostró un descenso más temprano ante la señal que simulaba el movimiento de la cabeza (Fig. 9, panel C, línea azul), donde en el ensayo número 700 comenzó un declive constante y pronunciado que le permitió alcanzar un valor de error de **0.47** en el ensayo 1300, para así continuar aprendiendo y mostrar un valor de **0.14** al terminar el entrenamiento. Posteriormente, el valor de error para la condición de mirada (Fig. 9, panel C, línea anaranjada) comenzó su descenso más pronunciado 400 ensayos después de la simulación de cabeza (i.e. ensayo 1100) alcanzando igualmente un valor de **0.47** hasta el ensayo número 1900 y manteniendo a partir de ahí un aprendizaje muy cercano al dado en la condición de cabeza, resultando en un error de **0.17** después de los 5000 ensayos de entrenamiento. Finalmente, la curva de error para la condición no social (Fig. 9, panel C, línea anaranjada) mostró un descenso incluso más demorado y gradual, comenzando con un valor de **0.73** que alcanza el **0.29** al final del entrenamiento, sin algún declive pronunciado.

En la red de **dirección** ante **percepción atípica** la disminución del error es nula para los tres estímulos durante prácticamente toda la primera mitad de la simulación; comienzan con un valor de error de **0.66** y durante los primeros 2000 ensayos se mantienen de la misma manera, lo que representa el 40% de la duración del entrenamiento. Llegando al ensayo 2100, la primera

curva que muestra un descenso es la del estímulo que simula una Cabeza direccionada (Fig. 9, panel D, línea azul); este descenso es gradual y termina por alcanzar un error de **0.29**. Una trayectoria similar se observa para la curva de la simulación de Mirada (Fig. 9, panel D, línea anaranjada), en donde el descenso del error comienza a ser perceptible en el ensayo 2700 para alcanzar un valor de **0.36** tras los 5000 ensayos. Finalmente, la curva de error para la simulación de un estímulo no social (Fig. 9, panel D, línea gris) presenta solo una ligera demora en comparación con la de Mirada, comenzando un declive en el nivel de error hasta el ensayo número 3100 y alcanzando un error de **0.40** al término del entrenamiento.

Teniendo en consideración las alteraciones presentes en TEA en procesos cognitivos como la atención y la percepción (Markram & Markram, 2010), nuestra simulación de percepción atípica diseñada a partir de la alteración de la pista para presentarla de manera ruidosa nos permite observar una trayectoria de desarrollo de la RJA demorada en comparación con la de Desarrollo Típico (DT) para las condiciones de movimiento y los distintos tipos de pista. No obstante, es evidente que, incluso ante las atipicidades perceptuales, la información dinámica permite una manifestación más temprana de RJA para los estímulos tipo cabeza, mirada y no social, siendo nuevamente el movimiento de cabeza la pista ante la que más pronto aprende la red a presentar una Respuesta a la Atención Conjunta.

### ***7.2.3. Hiperfocalización y percepción atípica (HF+PA)***

Las curvas de error de esta condición manifestaron tanto la atipicidad del comportamiento de la condición de hiperfocalización, como la demora de la condición de percepción atípica. De esta manera, el comportamiento de las curvas para esta simulación se caracteriza por ser demorado y atípico en la condición de movimiento y en la de dirección.

En la simulación **TEA\_MOV\_HF+PA**, el primer descenso de la curva de error comenzó en la pista de Mirada (Fig. 9, panel E, línea anaranjada) en el ensayo 900 con un valor de **0.67** y continuando durante el resto del entrenamiento con un declive bastante gradual y lento, alcanzando un valor de error de **0.29** en el ensayo 5000. Posteriormente, el comportamiento de las curvas de error para las pistas tipo Cabeza y estímulo No social (Fig. 9, panel E, líneas azul y gris, respectivamente) se comportan de una manera casi idéntica, salvo por mínimos momentos. Ambas comienzan con un valor de **0.74**, en el ensayo 1400 la pista No social tiene un valor de **0.67** mientras la de Cabeza se sitúa ligeramente por debajo en el mismo ensayo con un valor de **0.65**; a mitad del entrenamiento los valores muestran una inversión y la curva del estímulo no social se sitúa por debajo de la de Cabeza con valores de **0.47** y **0.51**, respectivamente; al final, las dos curvas muestran en el ensayo 5000 un valor de error de **0.25**.

Por otra parte, en la condición **TEA\_DIR\_HF+PA** fue en la que menos aprendizaje se alcanzó para los tres tipos de estímulos al final del entrenamiento de las redes, siendo la simulación de Mirada direccionada (Fig. 9, panel F, línea anaranjada) la que mostró primero un descenso en el error, manifestando inicialmente un valor de **0.65** que mantiene durante 1000 ensayos y luego disminuye hasta alcanzar un **0.42** en el ensayo 3000, número que se mantiene hasta el final del entrenamiento. Al igual que en la condición de movimiento para hiperfocalización y percepción atípica descrita en el párrafo anterior, la simulación de Cabeza y estímulo Co social (Fig. 9, panel F, líneas azul y gris, respectivamente) presentaron un comportamiento casi idéntico con un aprendizaje claramente lento y gradual durante la mayor parte del entrenamiento; no obstante la demora, en el ensayo 4500 la curva de error de la simulación de Cabeza comienza a tomar una ligera ventaja sobre la simulación de estímulo No social, logrando alcanzar el mismo valor de error que la simulación de Mirada al final del

entrenamiento y dejando al estímulo No social ligeramente por encima de ella con valores de **0.42** y **0.43**, respectivamente.

Una característica particularmente interesante del comportamiento de la simulación TEA con hiperfocalización y percepción atípica es que, tanto en la condición de movimiento como en la de dirección, el primer estímulo que parece favorecer el aprendizaje de la RJA es el que simula la mirada y, aunque durante la mayor parte del entrenamiento encabeza el aprendizaje con los valores de error más bajos, a lo largo de los últimos 1000 ensayos el descenso del error parece estancarse a tal punto que las curvas de los dos estímulos restantes (i.e. cabeza y no social) la sobrepasan (en el caso de movimiento) o se acercan a hacerlo (en el caso de dirección). Esta trayectoria parecería corresponder con los hallazgos de Klin y Jones (2013) al realizar un estudio de atención hacia el área de los ojos en infantes que posteriormente son diagnosticados con TEA y descubrir que, durante los primeros meses de vida, la atención hacia el área de los ojos en infantes que desarrollan TEA y que no lo desarrollan es la misma; sin embargo, entre los 2 y 6 meses, los(as) que se encontrarán en el espectro del autismo presentan un declive en esta conducta, manifestando lo que posteriormente se identifica como signo temprano del TEA: el pobre contacto visual.

Frente a esta situación, una vez más resulta importante cuestionar la importancia que se le ha dado en el quehacer científico y clínico al seguimiento de mirada como el gran facilitador del desarrollo de la Respuesta a la Atención Conjunta, sin reflexionar respecto a la posibilidad de su aprendizaje alentado y favorecido por el procesamiento de otro tipo de información que puede variar tanto en las características dinámicas como en su relevancia social, o dicho de manera apegada a la presente investigación: la cantidad de información disponible y su procesamiento.

## 8. CONCLUSIONES

En este trabajo se estudió la Respuesta a la Atención Conjunta (RJA) y su manifestación ante diferentes tipos de estímulos, incluyendo sociales: mirada y cabeza; y no sociales: líneas, con diferentes propiedades dinámicas: movimiento y dirección final. Esto se realizó a través de modelos computacionales que nos permitieron simular las características de procesamiento de población típica y población diversa con Trastorno del Espectro Autista (TEA).

Se encontró que, durante el desarrollo típico de la RJA, entre más información es percibida (Cabeza>Mirada>No social y Movimiento>Dirección), menos tiempo toma su desarrollo y manifestación, sugiriendo así que las diferencias en respuestas ante estos estímulos, que han sido tradicionalmente atribuidas a su carácter social, pueden resultar también de la cantidad de información que proveen. Así mismo, nuestras simulaciones describen una trayectoria típica de dos fases: la primera fase se centra en la detección y seguimiento de las pistas (proceso perceptual) y la segunda en la detección de un *target* con base en la lectura de las pistas (proceso asociado a coordinación social).

Considerando la diversidad de estilos perceptuales en la población con Trastorno del Espectro del Autismo, en este estudio se modelaron dos principales variables que han sido caracterizadas en TEA: la percepción atípica y la hiperfocalización. Se identificó una trayectoria *demorada* del desarrollo de la RJA en simulaciones de percepción atípica; por otro lado, se observó una trayectoria *atípica* en simulaciones de hiperfocalización; y finalmente, una trayectoria tanto *atípica* como *demorada* en simulaciones donde se combinaron las variables de hiperfocalización y percepción atípica. Con base en estos hallazgos, consideramos que, aunque la diversidad de descubrimientos y propuestas que se han hecho a partir de la investigación y análisis de la Respuesta a la Atención Conjunta en TEA podrían tener como base las diferencias

metodológicas de cada estudio, también es posible que las características del neurodesarrollo de cada individuo influyan en los resultados obtenidos, favoreciendo o limitando algunos estilos perceptuales, atencionales y de procesamiento. Una trayectoria atípica, demorada, o atípica y demorada, finalmente puede manifestar un mismo hito en diferentes momentos y a través de diferentes caminos.

Nuestro análisis de las propiedades de los estímulos desencadenantes de la Respuesta a la Atención Conjunta acentúa fuertemente la importancia de abordar el aprendizaje y desarrollo de la RJA como un proceso longitudinal que se beneficia de un tipo y cantidad de información variable a través del tiempo y con base en las habilidades que se van construyendo gracias al desarrollo neurobiológico. Dicho lo cual, plantear la existencia de “Estadios del Desarrollo de la Atención Conjunta” nos permitiría observar algunas diferencias y “contraposiciones” entre los resultados que se han propuesto hasta el momento, como manifestaciones de la RJA en momentos diferentes de su desarrollo y posiblemente mediadas por habilidades cognitivas diversas también.

La comprensión y estudio de la neurodiversidad demanda una mirada amplia, global y dinámica en la que sean considerados los antecedentes ontogenéticos y la propia trayectoria de desarrollo individual para potenciarla. Hallazgos como el que nos brinda esta investigación nos alienta a la reflexión sobre los métodos de estimulación de la Atención Conjunta en los abordajes terapéuticos y la manera de incrementar su eficacia a través del análisis y acción personalizada. Ya que la atención hacia los estímulos sociales y su procesamiento podría ser obstaculizada por la cantidad de información que ellos brindan y no precisamente por su carácter “social”, resulta interesante considerar un abordaje terapéutico enfocado en estimular la capacidad de

procesamiento de información como base para el desarrollo de habilidades más complejas como la atención conjunta y el seguimiento de mirada.

Finalmente, el presente abordaje vía modelamiento computacional representa un apoyo para las hipótesis siguientes, que podrán explorarse empíricamente en estudios futuros: 1) las señales visuales más informativas permiten conductas de atención conjunta más tempranas y eficaces, particularmente la Respuesta a la Atención Conjunta, aunque solo en desarrollo típico (y en TEA cuando solo se caracteriza con percepción atípica); 2) el movimiento percibido de estímulos visuales es el primer facilitador para el desarrollo de la RJA en población con desarrollo típico y en TEA, independientemente de los estilos de procesamiento; 3) en una trayectoria típica del desarrollo, la RJA se manifiesta con mayor facilidad (menor latencia y mayor acierto) ante estímulos sociales; y 4) el desarrollo de la RJA en TEA presenta trayectorias distintas a la del desarrollo típico, pudiendo caracterizarse como demorada, atípica o demorada y atípica. La atipicidad y/o demora de las trayectorias de desarrollo de la Respuesta a la Atención Conjunta ante los distintos estilos de procesamiento del Trastorno del Espectro Autista también es modulada por el tipo de estímulos; socialmente relevantes (más informativos) o no sociales (menos informativos).

Nuestro proyecto implica dos importantes limitaciones propias de la metodología aplicada. Por una parte, los modelos neurocomputacionales y su diseño no pueden abstraer completamente la realidad, por lo que solo poseen la información más relevante del fenómeno a investigar y podrían ignorar el impacto de otras variables no simuladas. Asimismo, la manipulación aislada y el control minucioso de las pistas informativas que la simulación de un episodio de atención conjunta nos permite es imposible de alcanzar en una situación naturalista, por lo que los resultados deberán siempre contrastarse y generalizarse con delicadeza.

En el futuro, la línea de investigación, metodología y resultados de nuestro estudio pueden permitir la comparación con estudios experimentales longitudinales para profundizar y ampliar la caracterización de los estímulos desencadenantes y/o facilitadores en la(s) trayectoria(s) del desarrollo de la Respuesta a la Atención Conjunta.

## 9. REFERENCIAS

- Adrien, J. L., Rossignol-Deletang, N., Martineau, J., Couturier, G., & Barthelemy, C. (2001). Regulation of cognitive activity and early communication development in young autistic, mentally retarded, and young normal children. *Developmental Psychobiology*, *39*(2), 124–136. <https://doi.org/10.1002/dev.1036>
- American Psychiatric Association. (2014). *DSM-5. Manual Diagnóstico y Estadístico de los Trastornos Mentales DSM-5*. <https://www.medicapanamericana.com/materialesComplementarios/DSM-5/DSM-5.aspx>
- Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television “portrait” is looking. *The American Journal of Psychology*, *82*(4), 474–489.
- Astor, K., Thiele, M., & Gredebäck, G. (2021). Gaze following emergence relies on both perceptual cues and social awareness. *Cognitive Development*, *60*, 101121. <https://doi.org/10.1016/j.cogdev.2021.101121>
- Baron-Cohen, S. (1995a). *Mindblindness: An essay on autism and theory of mind* (pp. xxii, 171). The MIT Press.
- Baron-Cohen, S. (1995b). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. *Joint Attention: Its Origins and Role in Development*.
- Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, *31*(9), 907–922. [https://doi.org/10.1016/0028-3932\(93\)90147-r](https://doi.org/10.1016/0028-3932(93)90147-r)

- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8(6), 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>
- Butterworth, G., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, 3(3), 253–272. <https://doi.org/10.1177/016502548000300303>
- Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9(1), 55–72. <https://doi.org/10.1111/j.2044-835X.1991.tb00862.x>
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), 176–176. <https://doi.org/10.2307/1166214>
- Casey, B. J., Gordon, C. T., Mannheim, G. B., & Rumsey, J. M. (1993). Dysfunctional attention in autistic savants. *Journal of Clinical and Experimental Neuropsychology*, 15(6), 933–946. <https://doi.org/10.1080/01688639308402609>
- Chawarska, K., Klin, A., & Volkmar, F. (2003). Automatic Attention Cueing Through Eye Movement in 2-Year-Old Children With Autism. *Child Development*, 74(4), 1108–1122. <https://doi.org/10.1111/1467-8624.00595>
- Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. En *Joint attention: Its origins and role in development* (pp. 61–83). Lawrence Erlbaum Associates, Inc.
- Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, 34(1), 28–38. <https://doi.org/10.1037/0012-1649.34.1.28>

- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176. <https://doi.org/10.1126/science.7375928>
- Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2018). The Developmental Origins of Gaze-Following in Human Infants. *Infancy*, *24*.  
<https://doi.org/10.1111/infa.12276>
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.  
<https://doi.org/10.1146/annurev.ne.18.030195.001205>
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117.  
<https://doi.org/10.1037/0096-3445.115.2.107>
- Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, *14*(2), 179–211.  
[https://doi.org/10.1207/s15516709cog1402\\_1](https://doi.org/10.1207/s15516709cog1402_1)
- Elsabbagh, M., & Johnson, M. H. (2016). Autism and the Social Brain: The First-Year Puzzle. *Biological Psychiatry*, *80*(2), 94–99. <https://doi.org/10.1016/j.biopsych.2016.02.019>
- Fantz, R. L. (1961). The Origin of Form Perception. *Scientific American*, *204*(5), 66–73.
- Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, *7*(6), 705–718.  
<https://doi.org/10.1080/13506280050144399>
- Fausey, C., Jayaraman, S., & Smith, L. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, *152*, 101–107. <https://doi.org/10.1016/j.cognition.2016.03.005>

- Garretson, H. B., Fein, D., & Waterhouse, L. (1990). Sustained attention in children with autism. *Journal of Autism and Developmental Disorders*, *20*(1), 101–114.  
<https://doi.org/10.1007/BF02206860>
- Gernsbacher, M. A., Stevenson, J. L., Khandakar, S., & Goldsmith, H. H. (2008). Why Does Joint Attention Look Atypical in Autism? *Child development perspectives*, *2*(1), 38–45.  
<https://doi.org/10.1111/j.1750-8606.2008.00039.x>
- Gregory, N. J., Hermens, F., Facey, R., & Hodgson, T. L. (2016). The developmental trajectory of attentional orienting to socio-biological cues. *Experimental Brain Research*, *234*(6), 1351–1362. <https://doi.org/10.1007/s00221-016-4627-3>
- Hainline, L. (1978). Developmental changes in visual scanning of face and nonface patterns by infants. *Journal of Experimental Child Psychology*, *25*(1), 90–115.  
[https://doi.org/10.1016/0022-0965\(78\)90041-3](https://doi.org/10.1016/0022-0965(78)90041-3)
- Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, *198*(4319), 853–855. <https://doi.org/10.1126/science.918670>
- Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*, *61*(1), 25–39.  
<https://doi.org/10.1016/j.bandc.2006.03.004>
- Harris, N. S., Courchesne, E., Townsend, J., Carper, R. A., & Lord, C. (1999). Neuroanatomic contributions to slowed orienting of attention in children with autism. *Cognitive Brain Research*, *8*(1), 61–71. [https://doi.org/10.1016/S0926-6410\(99\)00006-3](https://doi.org/10.1016/S0926-6410(99)00006-3)

- Haxby, J., Hoffman, E., & Gobbini, M. (2000). The Distributed Human Neural System for Face Perception. *Trends in cognitive sciences*, 4, 223–233. [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0)
- Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's Eyes Trigger Shifts of Visual Attention in Human Infants. *Psychological Science*, 9(2), 131–134. <https://doi.org/10.1111/1467-9280.00024>
- Houghton, G. (2005). *Connectionist Models in Cognitive Psychology*. Psychology Press.
- James, W. (1890). *The principles of psychology, Vol I*. (pp. xii, 697). Henry Holt and Co. <https://doi.org/10.1037/10538-000>
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, 6(10), 766–774. <https://doi.org/10.1038/nrn1766>
- Johnson, M. H., Posner, M., & Rothbart, M. (1991). Components of Visual Orienting in Early Infancy: Contingency Learning, Anticipatory Looking, and Disengaging. *Journal of cognitive neuroscience*, 3, 335–344. <https://doi.org/10.1162/jocn.1991.3.4.335>
- Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, 504(7480), 427–431. <https://doi.org/10.1038/nature12715>
- Karmiloff-Smith, A., Aschersleben, G., Schonon, S., Elsabbagh, A., Hohenberger, J., & Serres, J. (2010). Constraints on the Timing of Infant Cognitive Change: Domain-Specific or Domain-General? *European Journal of Developmental Science*, 4, 31–45. <https://doi.org/10.3233/DEV-2010-4103>
- Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., Fukuda, H., Kojima, S., & Nakamura, K. (1999). The human amygdala plays an important role in gaze

- monitoring: A PET study. *Brain*, 122(4), 779–783.  
<https://doi.org/10.1093/brain/122.4.779>
- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to nonsocial contingencies rather than biological motion. *Nature*, 459(7244), 257–261. <https://doi.org/10.1038/nature07868>
- Kylliäinen, A., & Hietanen, J. K. (2004). Attention orienting by another's gaze direction in children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 45(3), 435–444. <https://doi.org/10.1111/j.1469-7610.2004.00235.x>
- Leekam, S. R., López, B., & Moore, C. (2000). Attention and joint attention in preschool children with autism. *Developmental Psychology*, 36(2), 261–273.  
<https://doi.org/10.1037/0012-1649.36.2.261>
- Lempers, J. D. (1979). Young children's production and comprehension of nonverbal deictic behaviors. *The Journal of Genetic Psychology: Research and Theory on Human Development*, 135(1), 93–102. <https://doi.org/10.1080/00221325.1979.10533420>
- Lord, C. (2015). *ADOS 2, Escala de Observación para el Diagnóstico del Autismo*. TEA.
- Luria, A. (1984). *Atencion y Memoria*. Martínez Roca.
- Markram, K., & Markram, H. (2010). The Intense World Theory – A Unifying Theory of the Neurobiology of Autism. *Frontiers in Human Neuroscience*, 4.  
<https://www.frontiersin.org/article/10.3389/fnhum.2010.00224>
- Mayer, D. L., & Dobson, V. (1982). Visual acuity development in infants and young children, as assessed by operant preferential looking. *Vision Research*, 22(9), 1141–1151.  
[https://doi.org/10.1016/0042-6989\(82\)90079-7](https://doi.org/10.1016/0042-6989(82)90079-7)

- Michel, C., Kayhan, E., Pauen, S., & Hoehl, S. (2021). Effects of Reinforcement Learning on Gaze Following of Gaze and Head Direction in Early Infancy: An Interactive Eye-Tracking Study. *Child Development*, *n/a(n/a)*. <https://doi.org/10.1111/cdev.13497>
- Moore, C., Angelopoulos, M., & Bennett, P. (1997). The role of movement in the development of joint visual attention. *Infant Behavior & Development*, *20*(1), 83–92. [https://doi.org/10.1016/S0163-6383\(97\)90063-1](https://doi.org/10.1016/S0163-6383(97)90063-1)
- Munakata, Y., & McClelland, J. (2003). Connectionist Models of Development. *Developmental Science*, *6*, 413–429. <https://doi.org/10.1111/1467-7687.00296>
- Mundy, P. (2016). *Autism and joint attention: Development, neuroscience, and clinical fundamentals* (pp. xviii, 350). Guilford Press.
- Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., & Seibert, J. (2003). *EARLY SOCIAL COMMUNICATION SCALES (ESCS)*. 65.
- Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*, *24*(3), 247–257. <https://doi.org/10.1007/BF02172225>
- Perrett, D. I., & Mistlin, A. J. (1990). Perception of facial characteristics by monkeys. In *Comparative perception, Vol. 2: Complex signals* (pp. 187–215). John Wiley & Sons.
- Plunkett, K., & Elman, J. (1997). *Exercises in Rethinking Innateness: A Handbook for Connectionist Simulations*. A Bradford Book.
- Posner, M. I., & Petersen, S. E. (1989). The Attention System of the Human Brain. *Annals of Neuroscience*.

- Puce, A., Allison, T., Gore, J., & McCarthy, G. (1995). Puce, A., Allison, T., Gore, J. C. & McCarthy, G. Face-sensitive regions in human extrastriate cortex studied by functional MRI. *J. Neurophysiol.* 74, 1192-1199. *Journal of neurophysiology*, 74, 1192–1199.
- Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. *Neuroscience & Biobehavioral Reviews*, 32(1), 123–142. <https://doi.org/10.1016/j.neubiorev.2007.06.004>
- Rumelhart, D. E., McClelland, J. L., & Group, P. R. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations* (Vol. 1). A Bradford Book.
- Rutherford, M. D., & Krysko, K. M. (2008). Eye direction, not movement direction, predicts attention shifts in those with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38(10), 1958–1965. <https://doi.org/10.1007/s10803-008-0592-4>
- Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, 253(5489), 265–266. <https://doi.org/10.1038/253265a0>
- Schaffer, H. R. (1984). The child's entry into a social world. *Behavioural Development: A Series of Monographs*, 236–236.
- Senju, A., Tojo, Y., Dairoku, H., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry*, 45(3), 445–458. <https://doi.org/10.1111/j.1469-7610.2004.00236.x>
- Shiffrar, M., Lichtey, L., & Chatterjee, S. H. (1997). The perception of biological motion across apertures. *Perception & Psychophysics*, 59(1), 51–59. <https://doi.org/10.3758/BF03206847>

- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 809–813. <https://doi.org/10.1073/pnas.0707021105>
- Strube, G. (2001). Cognitive Modeling: Research Logic in Cognitive Science. En N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 2124–2128). Pergamon. <https://doi.org/10.1016/B0-08-043076-7/00588-X>
- Styles, E. A. (2010). *Psicología de la Atención*. Editorial Universitaria Ramón Areces.
- Swettenham, J., Condie, S., Campbell, R., Milne, E., & Coleman, M. (2003). Does the perception of moving eyes trigger reflexive visual orienting in autism? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1430), 325–334. <https://doi.org/10.1098/rstb.2002.1203>
- Tanaka, J., & Gauthier, I. (1997). Expertise in object and face recognition. En *Perceptual learning* (pp. 83–125). Academic Press.
- Thomas, M. S. C., & McClelland, J. L. (2001). Connectionist Models of Cognition. En R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (1a ed., pp. 23–58). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816772.005>
- Thomas, M. S. C., & McClelland, J. L. (2008). Connectionist Models of Cognition. En R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 23–58). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816772.005>
- Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. <https://doi.org/10.1016/j.jhevol.2006.10.001>

- Tovar, A. (en prensa). Capítulo 8. Modelos Neurocomputacionales. En *Introducción Práctica a los Modelos Cuantitativos de Cognición y Conducta*. (Federico Sanabria&Oscar Zamora (Eds.)). UNAM.
- Valenza, E., Otsuka, Y., Bulf, H., Ichikawa, H., Kanazawa, S., & Yamaguchi, M. (2015). Face Orientation and Motion Differently Affect the Deployment of Visual Attention in Newborns and 4-Month-Old Infants. *PLoS ONE*, *10*, e0136965.  
<https://doi.org/10.1371/journal.pone.0136965>
- Vecera, S. P., & Rizzo, M. (2006). Eye gaze does not produce reflexive shifts of attention: Evidence from frontal-lobe damage. *Neuropsychologia*, *44*(1), 150–159.  
<https://doi.org/10.1016/j.neuropsychologia.2005.04.010>
- Vlamings, P. H. J. M., Stauder, J. E. A., van Son, I. A. M., & Mottron, and L. (2005). Atypical Visual Orienting to Gaze- and Arrow-Cues in Adults with High Functioning Autism. *Journal of Autism and Developmental Disorders*, *35*(3), 267–277.  
<https://doi.org/10.1007/s10803-005-3289-y>
- Wainwright-Sharp, J. A., & Bryson, S. E. (1993). Visual orienting deficits in high-functioning people with autism. *Journal of Autism and Developmental Disorders*, *23*(1), 1–13.  
<https://doi.org/10.1007/BF01066415>
- Wang, Q., Celebi, F., Flink, L., Greco, G., Wall, C., Prince, E., Lansiquot, S., Chawarska, K., Kim, E., Boccanfuso, L., Dinicola, L., & Shic, F. (2015). *Interactive Eye Tracking for Gaze Strategy Modification*. <https://doi.org/10.1145/2771839.2771888>
- Wang, Q., Wall, C. A., Barney, E. C., Bradshaw, J. L., Macari, S. L., Chawarska, K., & Shic, F. (2020). Promoting social attention in 3-year-olds with ASD through gaze-contingent eye tracking. *Autism Research*, *13*(1), 61–73. <https://doi.org/10.1002/aur.2199>

- Yu, C., & Smith, L. B. (2013). Joint Attention without Gaze Following: Human Infants and Their Parents Coordinate Visual Attention to Objects through Eye-Hand Coordination. *PLOS ONE*, 8(11), e79659. <https://doi.org/10.1371/journal.pone.0079659>
- Yu, C., & Smith, L. B. (2017a). Multiple Sensory-Motor Pathways Lead to Coordinated Visual Attention. *Cognitive Science*, 41 Suppl 1, 5–31. <https://doi.org/10.1111/cogs.12366>
- Yu, C., & Smith, L. B. (2017b). Hand-Eye Coordination Predicts Joint Attention. *Child Development*, 88(6), 2060–2078. <https://doi.org/10.1111/cdev.12730>
- Zilbovicius, M., Meresse, I., Nadia, C., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in neurosciences*, 29, 359–366. <https://doi.org/10.1016/j.tins.2006.06.004>
- Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, 23(2–3), 143–152. <https://doi.org/10.1016/j.ijdevneu.2004.05.001>