



OPEN

## Feasibility of a 2-minute eye-tracking protocol to support the early identification of autism

Lacey Chetcuti<sup>1</sup>, Kandice J. Varcin<sup>2,3</sup>, Maryam Boutrus<sup>3</sup>, Jodie Smith<sup>1,4</sup>, Catherine A. Bent<sup>1</sup>, Andrew J. O. Whitehouse<sup>3</sup> & Kristelle Hudry<sup>1</sup>✉

We tested the potential for *Gazefinder* eye-tracking to support early autism identification, including feasible use with infants, and preliminary concurrent validity of trial-level gaze data against clinical assessment scores. We embedded the ~2-min 'Scene 1S4' protocol within a comprehensive clinical assessment for 54 consecutively-referred, clinically-indicated infants (prematurity-corrected age 9–14 months). Alongside % tracking rate as a broad indicator of feasible assessment/data capture, we report infant gaze data to pre-specified regions of interest (ROI) across four trial types and associations with scores on established clinical/behavioural tools. Most infants tolerated *Gazefinder* eye-tracking well, returning high overall % tracking rate. As a group, infants directed more gaze towards social vs. non-social (or more vs. less socially-salient) ROIs within trials. Behavioural autism features were correlated with increased gaze towards non-social/geometry (vs. social/people) scenes. No associations were found for gaze directed to ROIs within other stimulus types. Notably, there were no associations between developmental/cognitive ability or adaptive behaviour with gaze towards any ROI. *Gazefinder* assessment seems highly feasible with clinically-indicated infants, and the people vs. geometry stimuli show concurrent predictive validity for behavioural autism features. Aggregating data across the ~2-min autism identification protocol might plausibly offer greater utility than stimulus-level analysis alone.

Autism spectrum disorder (hereafter, *autism*) is diagnosed when behavioural impairments converge across two broad functional domains: social-communication and behavioural flexibility<sup>1</sup>. Individual presentation is otherwise highly heterogeneous<sup>2</sup> and identification and diagnosis rely on clinical judgement of behavioural presentation, typically through direct observation and informant report<sup>3</sup>. This process can be protracted, bringing substantial delays to young autistic children and their families accessing critical therapeutic supports. Hence, there is an identified need for more accurate and efficient assessment processes, particularly in identifying emerging signs of autism during infancy<sup>4</sup> to facilitate supports while the developing brain is highly plastic, and before possible associated disability becomes entrenched<sup>5</sup>.

Technology offers the potential to improve diagnostic accuracy and efficiency by reducing the current reliance on the experience and confidence of individual clinical professionals<sup>6,7</sup>. Technology also has the capacity to support the earlier identification of autism through biomarker discovery<sup>8</sup>. Eye-tracking is one promising such approach, offering objective insights into the emergence and consolidation of autism in early childhood. Indeed, the application of eye-tracking in research programs and autism clinics has been steadily increasing<sup>9</sup>.

### Eye-tracking as potential autism identification tool

The potential for eye-tracking to contribute to the earlier identification and diagnosis of autism is grounded in well-documented visual attention differences between people with and without autism. In a seminal meta-analysis, Frazier et al.<sup>10</sup> quantified the magnitude and significance of autism-related gaze differences, aggregating data across 122 independent case-control studies to show robust gaze differences across a wide range of stimulus types and specific within-stimulus regions of interest (ROIs). Notably, there were marked differences between autistic and non-autistic individuals in gaze towards non-social vs. social ROIs, and towards eye- and whole-face ROIs compared to nose-, mouth-, and body ROIs within human/interactive stimuli. Findings from another,

<sup>1</sup>Department of Psychology, Counselling and Therapy, School of Psychology and Public Health, La Trobe University, Melbourne, VIC, Australia. <sup>2</sup>Menzies Health Institute Queensland, Griffith University, Gold Coast, Australia. <sup>3</sup>Telethon Kids Institute, University of Western Australia, Perth, Australia. <sup>4</sup>School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, VIC, Australia. ✉email: k.hudry@latrobe.edu.au

more recent meta-analysis of 22 eye-tracking studies similarly concluded that gaze towards eye-regions of the human face and towards non-social scene ROIs effectively discriminated adult viewers with and without autism<sup>11</sup>.

Adjacent literature has evaluated the clinical utility of eye-tracking based visual social attention measurement to support autism diagnosis, including among young children deemed to be at increased likelihood due to family history of autism or the presentation of suggestive clinical/behavioural indicators (e.g.,<sup>12,13</sup>). Here, studies have provided complimentary evidence of increased number or duration of gaze fixations to non-social vs. social scenes (e.g., objects or geometric images vs. people<sup>14</sup>) and to typically less socially-salient elements within social scenes (e.g., people and faces) among 6-month-olds who later received an autism diagnosis<sup>15</sup>. Indeed, the performance properties of eye-tracking data for discriminating young children who do vs. do not subsequently receive an autism diagnosis is high, with sensitivity and specificity estimates similar to those from behavioural assessments historically adopted to support clinical decision-making<sup>16</sup>.

Despite the demonstrated potential of eye-tracking to discriminate individuals with and without autism, recent reviews have noted inconsistencies in individual study findings, partly attributable to methodological differences including features of the stimuli presented for viewing<sup>10,13</sup>. An important consideration towards achieving clinically-useful eye-tracking assessment is for standardized, validated and scalable paradigms. An approach whereby gaze data are gathered consistently—with standard apparatus, stimulus presentation, and pre-determined ROIs for gaze data capture—across samples from a given population of interest, might better support efforts to achieve an objective autism identification and diagnostic support tool, with stronger inferences drawn from a pool of studies founded on a common approach.

Additionally, to be clinically useful, eye-tracking technology will need to be feasibly incorporated within clinical care contexts. A promising approach will only be useful if the protocol is well-tolerated by its intended viewers—returning sufficient data to afford confident conclusions—and operable and interpretable by its intended users. In the context of seeking to support the earlier identification and diagnosis of autism, objective technology-based tools will only be useful if engaging for infant viewers and easily operable by clinical personnel with minimal technical training.

### The brief, standardised, and scalable GazeFinder protocol

Designed for ‘off the shelf’ use, JVCKENWOOD Corporation’s (JKC) *GazeFinder* eye-tracking technology (GP-100EA, detailed in Supplementary Figures S3–S6) comprises hardware and a standardized brief audio-visual presentation (referred to as ‘Scene 1S4’; Supplementary Figure S7, Tables S2–S3) featuring a collection of key stimuli previously used in autism eye-tracking research—five static and animated human face trials, six animated paired social and non-social scenes (i.e., people vs. geometry in various configurations), two biological motion PLDs animations, and two animated referential attention probes. Pre-specified ROIs (as detailed in Supplementary Figure S7, Tables S2–S3) allow for automatic visualisation of a viewer’s gaze data after assessment and ready exportation from the device for further in-depth analysis.

Previous research conducted independently from our team indicate that the *GazeFinder* ‘Scene 1S4’ animation effectively differentiates individuals with and without autism in childhood and adolescence/adulthood, with estimated sensitivity and specificity >74% in childhood (5–17 years)<sup>17</sup> and >80% in adolescence/adulthood<sup>18</sup>. Recent evidence further supports the feasibility of using *GazeFinder* with infants sampled from the general population, showing face validity with greater group mean-level gaze towards more vs. less socially-salient ROIs<sup>19</sup>.

### The current study

Extending investigation into the potential for *GazeFinder* as a standardised, scalable autism evaluation tool, we explored whether this technology might be feasibly incorporated into an assessment schedule around the time when first signs of autism often come to the attention of caregivers and professionals. Specifically, we tested the feasibility of using *GazeFinder* with a heterogeneous clinically-indicated sample of infants, enrolled into a larger research program due to identified early signs of autism. We incorporated *GazeFinder* GP-100EA devices with standardised ‘Scene 1S4’ autism assessment into appointments otherwise mirroring a clinical autism evaluation, for a proof-of-concept test of clinical implementation. Our two-fold focus was on: (1) the feasible incorporation of *GazeFinder* for quality data capture with infant viewers presenting with diverse clinical histories and presentation, and (2) the preliminary validity of gaze data to pre-specified ROIs evidenced through within-stimulus preferential attention and concurrent association with established clinical/behavioural assessments administered at the same appointment.

Regarding feasibility, we were interested in the successful incorporation of *GazeFinder* assessment within the broader clinical study protocol and tolerance by infant viewers. Specifically, we operationalised feasibility in terms of successful gaze detection during an initial calibration sequence, and the subsequent overall gaze data capture (i.e., % tracking rate) by *GazeFinder* across the ~2-min stimulus sequence.

Regarding preliminary validity, we hypothesised that infants as a group would show broad preferential attention towards social vs. non-social and more vs. less socially-salient ROIs within stimulus types: eyes (vs. mouth) of a human face (particularly in static presentations vs. animated scenes with mouth moving/talking), people (vs. geometry) in preferential attention pairs, upright (vs. inverted) biological motion point-light displays (PLD), and referential agent and target object (vs. distractors) in referential attention probes.

Further, we hypothesised that clinical assessment scores suggesting more signs of autism would be associated with more gaze towards non-social/less socially-salient ROIs, and less gaze towards social/more socially-salient ROIs. Given intended specificity of *GazeFinder*’s ‘Scene 1S4’ protocol to assess autism-related visual attention differences, we did not anticipate any significant associations between gaze data and clinical scores on broader developmental/cognitive ability or adaptive functioning assessments. However, given the important role of shared and referential (i.e., dyadic and triadic) attention for language development<sup>20</sup>, which is often delayed or different

in autism<sup>21</sup>, we anticipated that infants with poorer assessed language skills might show less gaze towards agent/target during specific referential attention scenes of the ‘Scene 1S4’ autism assessment.

## Methods

### Design and procedure

This evaluation was embedded within a larger research program enrolling community-referred infants showing early signs of possible emerging autism<sup>22,23</sup>, with experimental tasks (including eye-tracking) incorporated alongside a clinical/behavioural assessment battery. This research was approved by the Child and Adolescent Health Service Ethics Committee (CAHSEC Ref# 2016008EP; June 8th 2016) and performed in accordance with the Declaration of Helsinki with each caregiver providing written informed consent for their infant’s participation at the start of the appointment. This broader research program offered an opportune context within which to examine the feasible use of *Gazefinder* within an assessment schedule otherwise mirroring activities incorporated in clinical early identification and assessment for autism, and involving a participant sample with diverse clinical family medical histories and personal presentations. As *Gazefinder* devices were available for incorporation into the larger research program around recruitment mid-point, data reported here are for a subgroup (52%) of the full recruited cohort. Specifically, participants contributing data here were all consecutively-enrolled infants seen from January 2017 to February 2018.

Inclusion criteria for infant participants were prematurity-adjusted age 9–14-months (i.e., corrected if gestational age < 37 weeks), and clinical-indication of increased autism likelihood, operationalised as  $\geq 3$  of 5 key early autism behaviours according to the Social Attention and Communication Surveillance–Revised (SACS-R) developmental surveillance tool 12-month check<sup>24</sup>. Potentially eligible infants were referred from community Maternal and Child Health Nurses trained to undertake SACS-R monitoring with infants attending their routine child health/development checks, or following file review for infants engaged with a community Child Development Service<sup>22</sup>. In both cases, formal eligibility against the SACS-R tool was confirmed through discussion with each infant’s parent/caregiver at the point of study enrolment. The in-person assessment was scheduled where possible, within the following 2 weeks. Additional specified inclusion/exclusion criteria for the research were sufficient parental English-language capacity for full study participation, absence of any pre-identified developmental/neurological condition in the infant (including substantially pre-term birth < 32 weeks’ gestational age), and no family intention to relocate within the planned 2-year follow-up time.

Eligible infants whose parents consented to their participation attended an appointment comprising direct- and parent-reported clinical/behavioural assessments and experimental tasks (including *Gazefinder* eye-tracking once apparatus were available in January 2017). The clinical researchers responsible for completing all assessments had no specific disciplinary expertise of/extensive experience using eye-tracking technology, and were provided with an operational manual and one-off demonstration of *Gazefinder* use by its manufacturers. Clinical/behavioural measures were administered and scored blind to infants’ eye-tracking data (i.e., all analysis of *Gazefinder* data occurred after the analysis and reporting of clinical/behavioural data<sup>22</sup>).

### Clinical/behavioural phenotyping

The clinical/behavioural assessment protocol included the semi-structured, play-based Autism Observation Scale for Infants (AOSI)<sup>25</sup> offering a measure of emerging autism signs. Administration was by trained clinical researchers who showed high intra- and inter-rater agreement<sup>22,26</sup>. Higher AOSI total scores signal greater autism-related behaviours, with scores  $\geq 9$  suggestive of high likelihood of a later autism diagnosis<sup>27,28</sup>.

The clinical/behavioural protocol also included two standardised, norm-referenced measures of broader developmental/cognitive abilities, and adaptive behaviours that regularly feature in clinical autism research<sup>21,27,29</sup>. The Mullen Scales of Early Learning (MSEL) indexes overall developmental/cognitive abilities with excellent test–retest and inter-scoring reliability for infants aged  $\leq 24$ -months ( $r \geq 0.82$ )<sup>30</sup>. A summary Non-Verbal Developmental Quotient (NVDQ) was derived for analysis by averaging infants’ age-equivalence (AE) scores across Visual Reception and Fine Motor scales, dividing by infant chronological age (prematurity adjusted where relevant), and then multiplying the result by 100. Additionally, MSEL Receptive Language and Expressive Language scale AE scores were retained separately for analysis of specific association with emerging language skills. The Vineland Adaptive Behavior Scales-2nd Edition (VABS-II) is a caregiver-report adaptive behaviour measure, showing excellent test–retest and inter-rater reliability for infants aged  $\leq 24$ -months ( $r \geq 0.81$ )<sup>31</sup>. The VABS-II Adaptive Behavior Composite (ABC) Standard Scores (SS; population  $M = 100$ ,  $SD = 15$ ) was retained for analysis, which summarises functioning across Communication, Daily Living (i.e., personal independence), Socialization, and Motor skill domains.

Also of relevance within the larger trial protocol was the 191-item, caregiver-reported Infant Behavior Questionnaire-Revised (IBQ-R)<sup>32</sup> which quantifies temperament traits across three domains potentially informative of *Gazefinder* tolerability: activity level, distress to limitations, and duration of orienting. Drawing on data from the larger cohort, we have previously shown internal consistency reliability across the IBQ-R, and predictive validity of domain scores for later child behaviour problems<sup>33</sup>.

### Eye-tracking protocol

*Gazefinder* model GP-100 EA was used for the current study, with product specifications detailed in manufacturer Instruction Manual<sup>34,35</sup> and key details from these included in Supplementary Figures S3–6. This *Gazefinder* apparatus includes a 19-inch monitor (1280 × 1024 pixels) with integrated infrared light and camera to capture viewers’ corneal reflection of eye position at 50 Hz (i.e., 3000 samples/min). The ‘Scene 1S4’ autism assessment includes a sequence of static and animated stimuli (illustrated below and detailed in Supplementary Figure S7 and Tables S2–S3) developed by JKC. We confirm the permission from actors to appear in these images or that

images were purchased by JKC for this use (including in open-access publication), and that we have the permission of JKC to include the images presented here (and in Supplementary Materials) for illustrative purposes including in open-access publication.

#### Setup and calibration

While *Gazefinder* eye-tracking was typically conducted towards the end of each appointment after other clinical/behavioural assessments, staff could vary task presentation order to maximise infant engagement across the session. *Gazefinder* assessment was attempted with all consecutively-enrolled infants from the point the apparatus became available at the two study sites. Key features of the *Gazefinder* apparatus and recording environment setup are shown in Fig. 1 and Supplementary Figures S3–S6. Infants viewed *Gazefinder* whilst seated on a caregiver's lap. Caregivers were asked to close their eyes or avert their gaze during the stimulus presentation to avoid inadvertent adult vs. infant gaze data capture. Researchers used the *Gazefinder* on-screen visualisation to check correct infant positioning (also shown in Fig. 1) before triggering a 5-point gaze calibration check (centre- and four peripheral points) ahead of the main 'Scene 1S4' stimulus presentation.

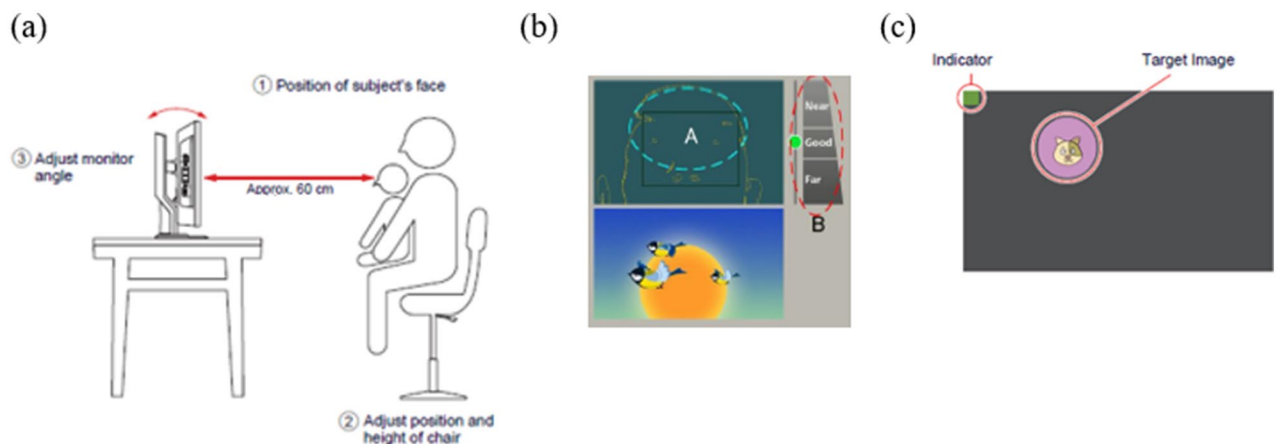
*Gazefinder* software was initially configured to begin 'Scene 1S4' presentation immediately following the calibration check sequence, without any opportunity for administering personnel to confirm calibration success. An examination of overall data capture (i.e., % tracking rate) for the first 24 infants assessed identified three instances of 0% tracking rate, suggesting undetected unsuccessful initial gaze calibration for these infants. The manufacturer made a minor software adjustment so that 'Scene 1S4' would only proceed to the main assessment if *Gazefinder* detected gaze to  $\geq 3$  (of 5) calibration points, including centre and at least two peripheral points (one top- and one bottom-of-screen). If not successful, *Gazefinder* would automatically repeat the calibration sequence (and only proceed to testing if gaze detected to  $\geq 3$  calibration points on this second attempt). If calibration was still unsuccessful, the assessment was automatically discontinued alerting the clinician to the issue. The clinician could then use their judgement about whether to reattempt assessment right away—via manual restart, including re-checking infant positioning using the on-screen visualisation, and engaging a new two-phase calibration attempt—or move to other tasks of the broader protocol before attempting *Gazefinder* eye-tracking again later in the session.

No maximum number of calibration attempts was specified, though the main animation sequence was only ever presented once per infant (i.e., if an infant's engagement seemed to wane following a successful initial calibration and commenced assessment, staff never discontinued the assessment mid-way to begin again later). Supplementary Figures S1–S2 outline the original and modified procedures for establishing gaze calibration before commencing the main assessment.

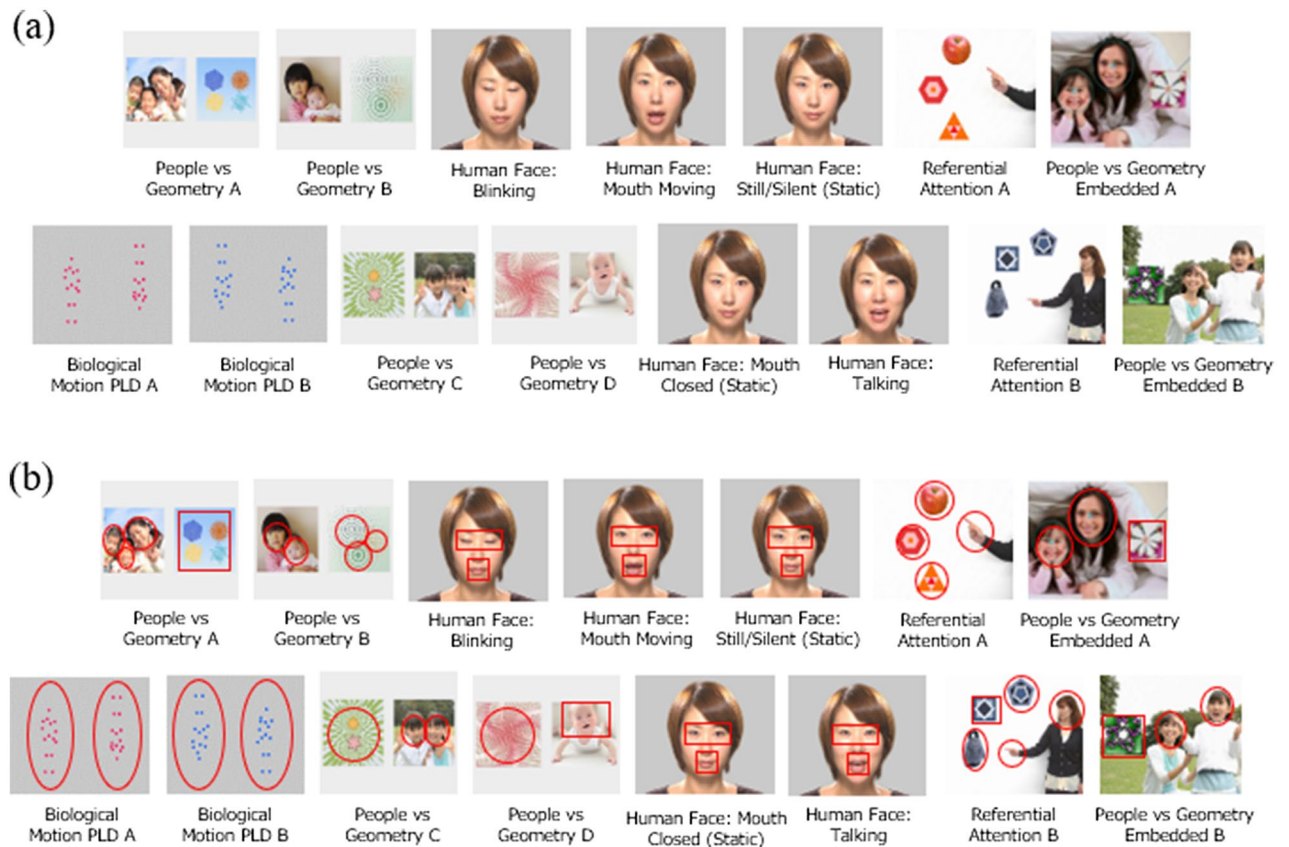
#### 'Scene 1S4' autism assessment sequence

Each infant viewed the identical 'Scene 1S4' animation (shown in Fig. 2; detailed in Supplementary Table S3); the same stimulus sequence used in previous studies with older child and adolescent/adult participants<sup>18,36</sup>. This ~2-min sequence includes short animations across the following four stimulus types, with *Gazefinder* software automatically summarising individual's gaze fixations to pre-specified ROIs:

- *Eyes vs. mouth* of a human face: one animated trial with eyes blinking (5 s), one animated trial with mouth opening and closing silently (2 s), two static trials (5 and 4 s, respectively), and one animated trial including speech (actress says [spoken in Japanese] "Hello. What is your name? "Let's play together"; 7 s);
- *People vs. geometry*: four trials with same-sized stimuli (range 4.7–5 s per trial; counterbalanced for left vs. right side of screen), and two trials with smaller geometric patterns embedded within larger social scenes (8 s per trial; also counterbalanced for side of screen);



**Figure 1.** (a) Guidance from *Gazefinder* Instruction Manual for setup, (b) on-screen visualisation to aide proper positioning of viewer in front of screen, and (c) example of attention-grabbing image for calibration check procedure.



**Figure 2.** (a) Key trials within *Gazefinder* ‘Scene 1S4’ autism assessment sequence, in presentation order (interspersed with attention-grabbers, not shown). (b) Pre-specified ROIs within each trial, representing paired social vs. non-social scenes or more vs. less socially-salient elements.

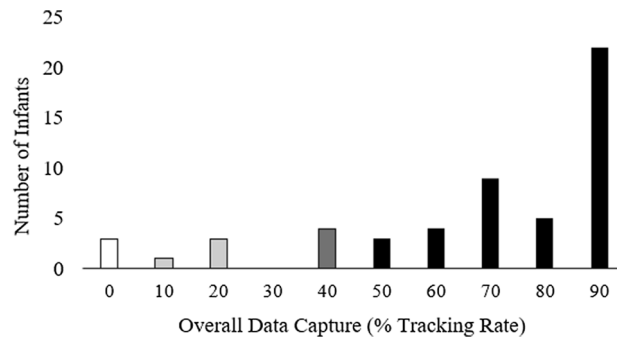
- Upright vs. inverted biological motion point-light type display (PLD): two trials (counterbalanced for left vs. right side of screen; 5–6 s per trial) with animation accompanied by instrumental song; and
- Referential attention probes: two trials each including a *target*, two *distractor* objects, and a referential *agent*; one trial showed only the pointing hand of the referential agent, while the other showed the full agent torso; 4 s per trial; and accompanied by “What is this?”).

Key trials were preceded by an attention-grabbing central stimulus—a cartoon animal motioning and saying (in Japanese) “Look, look” (1 s per trial; not shown in Fig. 2 and not contributing to gaze data extracted for analysis). Detailed parameters of trial timing and ROIs positioning are provided in Supplementary Tables S2–S3.

### Data analyses

*Gazefinder* software automatically summarised individual infants’ gaze data to the pre-specified ROIs for simple extraction of an encrypted, password-protected .csv file to collate with the other clinical/behavioural. Key *Gazefinder* data of interest here were % Tracking Rate—reflecting the overall amount of any gaze data captured as a proportion of ~2-min ‘Scene 1S4’ sequence—within stimulus % tracking rate, as well as gaze fixation durations to pre-specified ROIs within each of the four stimulus types (i.e., averaged across trials where multiple of the same stimulus type).

Associations between child clinical/behavioural characteristics and the % tracking rate (which was negatively skewed; as shown in Fig. 3) were examined using Spearman correlations. Relative gaze to ROI pairs within each stimulus type (i.e., social vs. non-social ROIs, or more- vs. less socially salient ROIs) was examined using paired-samples *t*-tests. Pearson/Spearman partial correlations—controlling for within-stimulus % tracking rate—were used to explore associations between child clinical/behavioural characteristics and (a) gaze towards each pre-specified ROI, and (b) ratio scores reflecting the *relative* extent of gaze towards social vs. non-social and more vs. less socially-salient ROIs (e.g., eye-vs.-mouth ratio; people-vs.-geometry ratio, etc.). We set no minimum threshold for overall or within-trial % tracking rate, retaining all participants with *any* available data for analyses. We set  $\alpha = 0.05$  and applied a false discovery rate (FDR) correction for multiple testing of gaze and clinical assessment association data.



**Figure 3.** Distribution of overall data capture across ~2-min *Gazefinder* “Scene 1S4” stimulus presentation.

## Results

### Participant characteristics

The sample of infants (74% male) was aged 9–16-months chronologically (and 9–14-months after correction for gestation < 37 weeks [ $n = 4$  infants]). Each infant was referred to the larger study by a community health practitioner on the basis of showing three- (26%), four- (26%), or five (48%) of five SACS-R eligibility markers suggesting possible emerging autism<sup>24</sup>. Table 1 presents group-level descriptive clinical/behavioural phenotyping data, and correlations among these measures. Group mean scores suggested moderate-to-high assessed autism behaviours (AOSI total score), broadly age-appropriate non-verbal developmental/cognitive abilities (NVDQ), language abilities delayed by ~ 3 months (receptive and expressive AE scores), and somewhat delayed adaptive skills (VABS ABC). Caregivers of a small subgroup of 10 infants (18.5%) reported that their infant had an older autistic sibling. Post-hoc review of caregiver-reported medical history notes revealed one infant with visual impairment warranting future surgical correction. We nevertheless retained this infant’s data, given *Gazefinder* assessment had been attempted/completed, and returned % Tracking Rate comparable to other infants suggesting no issues with gaze detection or sampling.

### Eye-tracking feasibility

Figure 3 shows the distribution of overall % tracking rate (range 0–99%) for the 54 infants with whom *Gazefinder* assessment was attempted. As already noted, three of the first 14 assessed returned 0% tracking rate, determined as having been due to examiner error with infant positioning relative to the monitor, resulting in failed calibration that was not detectable prior to proceeding with the main assessment (per original software programming; see Figure S1). Following additional manufacturer guidance on apparatus use and positioning (per Fig. 1 above) and minor software update to abort *Gazefinder* assessment if successful detection of gaze to  $\geq 3$  calibration points did not occur (see Figure S2), there were no further instances of 0% tracking Rate. Analysable data were therefore obtained for 51 of the 54 infants who completed *Gazefinder* assessment.

Among this final sample of 51 infants with analysable gaze data, overall % tracking rates exceeded 80% for over half of the sample (overall  $M = 77\%$ ,  $SD = 22.6\%$ ). Four infants returned minimal data (12–27% tracking rates); three infants were noticeably active/unsettled during their assessment (prompting early discontinuation of *Gazefinder* assessment for one due to fussiness, without re-attempt), and the fourth infant appeared attentive suggesting a potential issue with initial calibration/gaze detection. Excluding these four cases of < 30% tracking

	N	M (SD)	Range	AOSI	MSEL		VABS	
				Total score	NVDQ	Receptive	Expressive	ABC SS
Chronological Age	54	12.9 (1.83)	9–16	0.02, $p = 0.888$	– 0.29, $p < 0.05$	0.22, $p = 0.119$	0.28, $p < 0.05$	– 0.01, $p = 0.947$
AOSI								
Total score	54	9.50 (4.82)	2–28		– 0.14, $p = 0.329$	– 0.38, $p = 0.005$	– 0.27, $p < 0.05$	– 0.40, $p < 0.01$
MSEL								
NVDQ	53	105.9 (19.7)	61–154			0.43, $p = 0.002$	0.22, $p = 0.120$	0.02, $p = 0.892$
Receptive AE	54	9.9 (3.12)	2–17				0.44, $p < 0.001$	0.27, $p = 0.064$
Expressive AE	54	9.2 (3.02)	4–15					0.40, $p < 0.01$
VABS								
ABC SS	48	83.0 (9.9)	63–103					

**Table 1.** Summary of participant clinical/behavioural characterisation and associations among measures. AOSI Autism Observation Scale for Infants, MSEL Mullen Scales of Early Learning, NVDQ Non-Verbal Developmental Quotient, AE Age Equivalence, VABS Vineland Adaptive Behavior Scales, ABC Adaptive Behaviour Composite, SS Standard Score.

rate and reanalysing the data yielded substantively similar results (see Supplementary Tables S4–S5). Therefore, we present findings based on the full sample of 51 infants.

Associations between % tracking rate and child age, clinical/behavioural assessment scores, and caregiver-reported temperament measures are shown in Table 2. One single statistically significant association presented with IBQ-R Activity Level, suggesting lower % tracking rate for infants rated by caregivers as being more temperamentally active.

### Preliminary validity

#### *Within-trial preferential attention*

Figure 4 and Table 3 summarise infant gaze data towards pre-specified ROIs within each of five human face trials, and averaged across sets of four same-size people vs. geometry and two geometry-embedded trials, two biological motion PLD trials, and two referential attention probes. As a group, infants showed preferential attention towards eye (vs. mouth) ROIs of the human face during an initial blinking trial and a subsequent static trial, no preferential attention towards either ROI during a second static trial (presented later, separated from the first static trial by other stimulus types; see Fig. 2), and preferential attention towards mouth (vs. eyes) ROIs during two animated trials with mouth moving (silently and whilst talking).

As a group, infants also showed preferential attention towards people vs. geometry ROIs across both same-size and geometry-embedded trial types, and towards upright vs. inverted ROIs within biological motion PLD trials. During referential attention probes, infants showed more gaze towards target vs. distractor object ROIs, but most gaze towards referential agent ROIs (pointing hand, and face when visible).

#### *Associations of gaze and emerging autism behaviours*

Table 4 shows correlations between infant's scores on a measure of emerging autism signs (AOSI total scores) and gaze data to the pre-specified social and non-social or more vs. less socially-salient ROIs. AOSI total scores were significantly and moderately strongly correlated with gaze towards non-social (i.e., geometry) ROIs both within same-size ( $r=0.50$ ) and geometry-embedded trial types ( $r=0.44$ ). AOSI total scores were also moderately strongly correlated with gaze towards social (i.e., person) ROIs within these same trials, but this was not statistically significantly following FDR correction for multiple comparisons. There was no significant association of AOSI total scores with gaze towards ROIs in human face, biological motion, or referential attention trials.

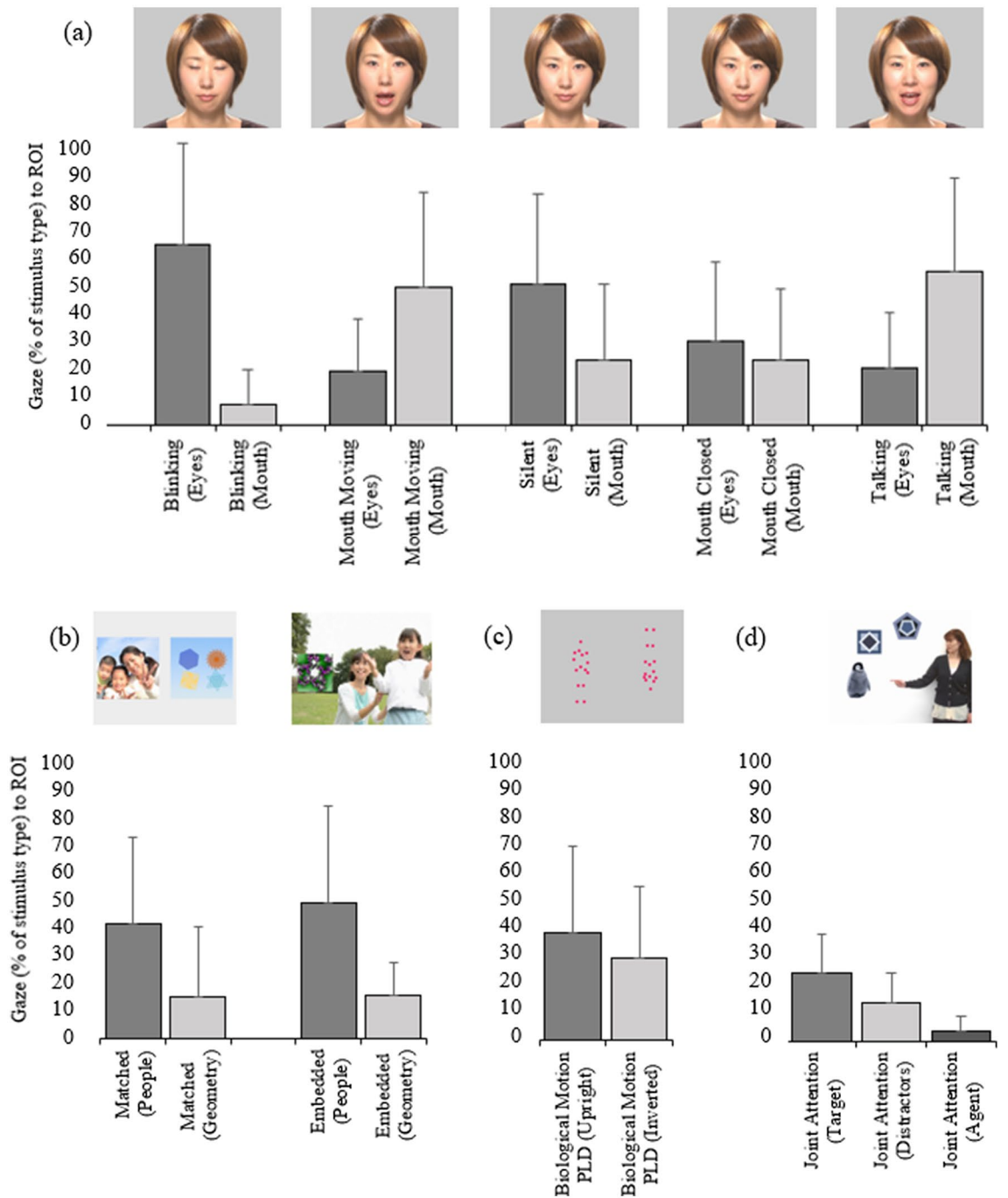
Table 5 shows additional correlations between infants' AOSI total scores and gaze data computed as proportionate measures (i.e., to people vs. geometry ROIs; to eyes vs. mouth ROIs of human face, etc.). Again, AOSI total scores were significantly and moderately strongly correlated with the proportionate index of gaze to people vs. geometry, across both matched size ( $r=-0.40$ ) and geometry-embedded trials ( $r=-0.41$ ). This was such that infants showing more signs of autism directed more gaze to concurrently presented geometry *relative* to people. Again, there were no significant correlations between AOSI Total Scores and proportionate gaze measures (to social vs. non-social, or more vs. less socially-salient ROIs) in human face, biological motion, or referential trials.

#### *Associations of gaze and other clinical phenotyping measures*

Also included in Tables 4 and 5 are correlations between measures of infants' broader developmental/cognitive abilities (MSEL NVDQ and Receptive and Expressive Language AE scores) and adaptive skills (VABS ABC SS) and gaze to pre-specified ROIs and computed proportionate ROI measures. Neither NVDQ nor ABC SS was significantly correlated with gaze data towards any ROI within any stimulus type. There was a suggested moderate-to-strong association between Receptive Language AE and gaze towards the target object ROI during

	N	Data capture
Chronological age	51	$r=0.18, p=0.216$
AOSI		
Total score	51	$r=0.12, p=0.401$
MSEL		
Non-verbal DQ	50	$r=0.09, p=0.540$
Receptive AE	51	$r=0.09, p=0.545$
Expressive AE	51	$r=0.13, p=0.353$
VABS		
Adaptive behaviour composite SS	48	$r=-0.19, p=0.216$
IBQ		
Activity level	49	$r=-0.30, p=0.036$
Duration of orienting	49	$r=0.11, p=0.437$
Distress to limitations	49	$r=0.02, p=0.869$

**Table 2.** Association of participant clinical/behavioural characterisation with overall rates of data capture. AOSI = Autism Observation Scale for Infants; MSEL = Mullen Scales of Early Learning; NVDQ = Non-Verbal Developmental Quotient; AE = Age Equivalence; VABS = Vineland Adaptive Behaviour Scales; SS = Standard Score; IBQ = Infant Behaviour Questionnaire.



**Figure 4.** Group mean-level preferential attention ( $\pm$  standard deviation) towards pre-specified social vs. non-social or more vs. less socially-salient ROIs across, four key trial types: (a) 5 presentations of human face (varying in static/animated elements); (b) 6 scenes pairing people vs. geometry (4 same-size, 2 with geometry embedded within larger social scene); (c) 2 biological motion PLD trials; and (d) 2 referential attention trials (one showing agent hand only; one showing full torso).

the referential attention probes that was statistically non-significant, however, with correction for multiple comparisons.

### Discussion

Autism diagnosis commonly occurs between age 3 and 5 years<sup>37,38</sup>, but shows stability from 24-months<sup>39</sup>. Neural indicators of emerging autism can be detected in early life<sup>40</sup> ahead of overt behavioural signs which often appear by the first birthday<sup>21</sup>. A protracted assessment and monitoring process is common across toddlerhood and into early childhood, however, with clinical judgement requiring the confident integration of information across directly observed and informant-reported behaviour, whilst taking into account developmental context and



	Group mean level gaze to ROIs		Difference
	Eyes	Mouth	
Human face			
Trial 1: blinking (silent)	0.60 (0.34) 0.00–1.00	0.07 (0.11) 0.00–0.38	$t(50) = 10.14, p < 0.001, d = 1.59$
Trial 2: mouth moving (silent)	0.18 (0.18) 0.00–0.82	0.46 (0.32) 0.00–0.89	$t(50) = 5.79, p < 0.001, d = 0.82$
Trial 3: static	0.47 (0.30) 0.00–1.00	0.22 (0.25) 0.00–0.88	$t(50) = 3.85, p < 0.001, d = 0.80$
Trial 4: static	0.28 (0.27) 0.00–0.97	0.22 (0.24) 0.00–0.88	$t(50) = 1.17, p = 0.248, d = 0.17$
Trial 5: talking	0.19 (0.18) 0.00–0.80	0.51 (0.31) 0.00–1.00	$t(50) = 5.75, p < 0.001, d = 0.99$
People vs. geometry	People	Geometry	
Same-size (4 trials averaged)	0.40 (0.21) 0.00–0.77	0.14 (0.12) 0.00–0.49	$t(50) = 7.07, p < 0.001, d = 1.22$
Embedded (2 trials averaged)	0.48 (0.29) 0.00–0.94	0.15 (0.18) 0.00–0.75	$t(50) = 6.07, p < 0.001, d = 1.00$
Biological motion PLD	Upright Figure	Inverted Figure	
(2 trials averaged)	0.38 (0.19) 0.00–0.79	0.29 (0.19) 0.00–0.62	$t(50) = 2.71, p = 0.009, d = 0.39$
Referential attention	Target	Distractor	
(2 trials)	0.13 (0.10) 0.00–0.38	0.04 (0.05) 0.00–0.22	$t(50) = 5.99, p < 0.001, d = 0.82$
	Agent face/hand		
	0.23 (0.14) 0.00–0.48	(vs. target)	$t(50) = 5.01, p < 0.001, d = 0.68$

**Table 3.** Summary of gaze towards ROIs across four trial types and between-region significance test. ROIs regions of interest.

known phenotypic heterogeneity<sup>4</sup>. Eye-tracking technology could expedite clinical decision-making by offering objective indication of autism-associated differences in visual social attention.

Some eye-tracking technologies require substantial user training, involve the time-consuming administration of stimuli across multiple trials, types, and/or paradigms, and can be relatively invasive if involving the use of goggles or head-mounted displays. All of these are important barriers to scalable clinical use. *Gazefinder* is a non-invasive eye-tracking system designed for easy off-the-shelf use by non-technical personnel. Its ‘Scene 1S4’ animation sequence offers a standardised, brief autism assessment protocol capturing gaze to pre-specified ROIs—social vs. non-social scenes, or more vs. less socially-salient within-scene elements—across four key stimulus types: trials showing a static or animated human face, trials tapping preferential attention to people vs. geometry, biological motion PLDs, and referential attention probes. Previous research has shown the potential for *Gazefinder* to support clinical autism decision-making, differentiating autistic from non-autistic individuals in childhood and adolescence/adulthood<sup>17,19,36</sup>. We tested the feasible use of *Gazefinder* with infants referred into a larger research program on the basis of showing signs of emerging autism, integrating the system within an appointment otherwise mirroring an age-appropriate clinical/behavioural autism assessment, conducted by clinical researchers without any particular prior technical eye-tracking expertise. Further, we appraised the preliminary validity of *Gazefinder* data against established and norm-referenced clinical/behavioural characterisation measures, quantifying signs of possible emerging autism, developmental/cognitive and adaptive behaviour skills, in early development.

We hypothesized that infants, as a group, would show broad preferential attention towards social vs. non-social and more vs. less socially-salient within-stimulus ROIs, but that individuals showing more behavioural signs of autism would show relatively greater gaze towards non-social scenes (i.e., geometry [vs. people] in paired preferential attention trials) and less socially-salient elements of scenes (i.e., mouth [vs. eyes] of human face in static and some animated trials; inverted [vs. upright] biological motion in PLD trials; and distractor objects [vs. target object and agent] in referential attention trials). We did not anticipate associations for gaze data with developmental/cognitive ability or adaptive skills measures, but did anticipate that receptive and expressive language skills might associate with gaze towards elements of referential attention trials reflecting shared/joint attention.

### Feasible integration within clinical autism assessment

With minimal training, non-technical clinical researchers feasibly incorporated *Gazefinder* within the broader clinical/behavioural assessment protocol for 54 consecutively-referred infants attending appointments from the time the system became available for use at two study sites. A small number of early failed assessments returning no data despite evident infant attentiveness (i.e., 0% Tracking Rate identified at assessment completion) indicated a potential systematic issue highlighting the importance of proper infant positioning relative to the screen, and benefit of a minor software adjustment that was made, following which there was no further experience of failed assessment.

Overall data capture (i.e., % tracking rate) was generally very high, over 80% for half of the sample; similar to previously-reported *Gazefinder* data capture tested with 4-to 11-month-olds infants sampled from the general population (sample mean 81% tracking rate<sup>19</sup>), and well above rates reported elsewhere in studies with infants using other eye-tracking systems<sup>41</sup>. Only four of the 51 successfully assessed infants in the current study returned < 30% overall tracking rates, with staff notes suggesting these could mostly be explained by waning inattention/increased distractibility during the stimulus presentation. Indeed, across the sample, among various child characteristics examined, only parent-reported temperamental activity level showed any association with % tracking rate. We therefore conclude that *Gazefinder* eye-tracking was generally feasibly implemented

ROI	AOSI		MSEL		VABS		MSEL			
	Total score		NVDQ		ABC SS		Receptive AE		Expressive AE	
Human face										
Blinking										
Eyes	$r=0.16$	$p=0.539$	$r_s=-0.06$	$p=0.964$	$r=0.26$	$p=0.410$	$r_s=-0.09$	$p=0.695$	$r=0.08$	$p=0.712$
Mouth	$r_s=-0.19$	$p=0.327$	$r_s=0.08$	$p=0.943$	$r_s=-0.14$	$p=0.838$	$r_s=0.12$	$p=0.771$	$r_s=0.16$	$p=0.576$
Mouth moving										
Eyes	$r_s=0.10$	$p=0.653$	$r_s=-0.15$	$p=0.943$	$r_s=0.04$	$p=0.881$	$r_s=-0.23$	$p=0.580$	$r_s=0.03$	$p=0.917$
Mouth	$r_s=-0.22$	$p=0.286$	$r_s=0.17$	$p=0.943$	$r_s=-0.17$	$p=0.838$	$r_s=0.08$	$p=0.771$	$r_s=-0.02$	$p=0.936$
Silent/still (static)										
Eyes	$r=0.16$	$p=0.539$	$r=-0.13$	$p=0.852$	$r=0.22$	$p=0.443$	$r_s=-0.21$	$p=0.300$	$r=-0.12$	$p=0.633$
Mouth	$r_s=-0.03$	$p=0.861$	$r_s=0.06$	$p=0.943$	$r_s=-0.16$	$p=0.838$	$r_s=0.28$	$p=0.580$	$r_s=0.19$	$p=0.563$
Mouth closed (static)										
Eyes	$r=0.04$	$p=0.768$	$r=0.11$	$p=0.852$	$r=0.20$	$p=0.443$	$r_s=0.01$	$p=0.955$	$r_s=0.20$	$p=0.633$
Mouth	$r_s=0.17$	$p=0.400$	$r_s=-0.14$	$p=0.943$	$r_s=-0.20$	$p=0.838$	$r_s=-0.08$	$p=0.771$	$r_s=-0.08$	$p=0.749$
Talking										
Eyes	$r_s=0.09$	$p=0.653$	$r_s=-0.17$	$p=0.943$	$r_s=0.13$	$p=0.838$	$r_s=-0.12$	$p=0.771$	$r_s=-0.09$	$p=0.749$
Mouth	$r=-0.04$	$p=0.768$	$r=0.01$	$p=0.965$	$r=0.00$	$p=0.991$	$r=0.13$	$p=0.612$	$r=0.09$	$p=0.712$
People vs geometry										
Same-size										
People	$r=-0.37$	$p=0.070$	$r=-0.01$	$p=0.965$	$r=0.09$	$p=0.626$	$r=0.34$	$p=0.070$	$r=0.16$	$p=0.633$
Geometry	$r_s=0.50$	$p<0.05$	$r_s=-0.03$	$p=0.943$	$r_s=-0.06$	$p=0.881$	$r_s=-0.27$	$p=0.580$	$r_s=-0.13$	$p=0.668$
Embedded										
People	$r_s=-0.35$	$p=0.056$	$r_s=-0.01$	$p=0.956$	$r_s=0.11$	$p=0.838$	$r_s=0.18$	$p=0.643$	$r_s=0.08$	$p=0.749$
Geometry	$r_s=0.44$	$p=0.05$	$r_s=0.10$	$p=0.943$	$r_s=-0.29$	$p=0.838$	$r_s=-0.12$	$p=0.771$	$r_s=-0.24$	$p=0.485$
Biological motion PLDs										
Upright figure	$r=0.13$	$p=0.539$	$r=-0.27$	$p=0.457$	$r=0.09$	$p=0.626$	$r_s=-0.21$	$p=0.300$	$r=0.04$	$p=0.862$
Inverted figure	$r=-0.14$	$p=0.539$	$r=0.02$	$p=0.965$	$r=-0.28$	$p=0.410$	$r=0.09$	$p=0.695$	$r=-0.15$	$p=0.633$
Referential attention										
Agent (face/hand)	$r=-0.07$	$p=0.754$	$r=-0.22$	$p=0.457$	$r=0.18$	$p=0.443$	$r_s=-0.02$	$p=0.955$	$r=-0.13$	$p=0.633$
Target	$r=-0.12$	$p=0.539$	$r=0.21$	$p=0.457$	$r=-0.17$	$p=0.443$	$r=0.34$	$p=0.070$	$r=0.35$	$p=0.110$
Distractors	$r_s=0.25$	$p=0.205$	$r_s=-0.06$	$p=0.943$	$r_s=0.05$	$p=0.881$	$r_s=-0.02$	$p=0.915$	$r_s=-0.22$	$p=0.556$

**Table 4.** Associations between gaze data to ROIs and scores on clinical/behavioural phenotyping measures for 51 infants with any available gaze tracking data. Significant values are in [bold]. All correlations are partial (controlling for within-stimulus % tracking rate) and  $p$  values are FDR corrected.  $r$  Pearson correlation,  $r_s$  Spearman's correlation, ROI region of interest, AOSI Autism Observation Scale for Infants, MSEL Mullen Scales of Early Learning, NVDQ Non-Verbal Developmental Quotient, VABS Vineland Adaptive Behavior Scales, ABC adaptive behaviour composite, AE age equivalence.

by clinical research staff and well-tolerated by infants with identified signs of possible emerging autism, referred around the first birthday.

### Some preliminary validity evidence for infant assessment

Evidence of group mean-level preferential attention was as expected. There were strong preferential effects for social vs. non-social scenes (i.e., people vs. geometry in paired scenes, whether matched size or geometry embedded within larger social scenes) and for more vs. less socially-salient within-scene elements within human face trials (i.e., eyes vs. mouth in an initial static trial and an animated blinking trial; but towards mouth vs. eyes in animated trials with mouth moving silently and/or when talking) and referential attention trials (i.e., most attention directed towards referential agent, followed by more attention towards target vs. distractor objects). There was also significantly greater gaze, at infant group mean-level gaze, towards upright vs. inverted biological motion PLDs animations, with small-to-moderate effect size. These data further evidence for the validity of the key stimulus types included in the 'Scene 1S4' *Gazefinder* autism assessment animation, consistent with previous accounts from use of this system with other samples/populations<sup>17–19</sup>, and from these paradigms presented to infants using other eye-tracking technologies<sup>10,11</sup>.

The current data offered more limited evidence of concurrent validity in the observed associations between infant gaze data and AOSI scores as an established measure of early behavioural autism signs. Infants expressing more autism behaviours directed more gaze to geometry ROIs in paired scenes, including both when computed as raw gaze towards the non-social element and a proportionate metric of gaze relative to the people ROIs. There was no association of autism behaviours, however, with gaze specifically towards social ROIs (i.e., people) within

	ROI	AOSI		MSEL		VABS		MSEL			
		Total score		NVDQ		ABC SS		Receptive AE		Expressive AE	
Human face											
Blinking											
Eyes vs mouth		$r_s = 0.20$	$p = 0.311$	$r_s = -0.11$	$p = 0.943$	$r_s = 0.14$	$p = 0.838$	$r_s = -0.10$	$p = 0.771$	$r_s = -0.16$	$p = 0.576$
Mouth moving											
Eyes vs mouth		$r_s = 0.23$	$p = 0.280$	$r_s = -0.15$	$p = 0.943$	$r_s = 0.08$	$p = 0.881$	$r_s = -0.12$	$p = 0.771$	$r_s = -0.05$	$p = 0.887$
Silent/still (static)											
Eyes vs mouth		$r_s = 0.04$	$p = 0.846$	$r_s = -0.04$	$p = 0.943$	$r_s = 0.21$	$p = 0.838$	$r_s = -0.25$	$p = 0.580$	$r_s = -0.13$	$p = 0.668$
Mouth closed (static)											
Eyes vs mouth		$r_s = -0.10$	$p = 0.653$	$r_s = -0.05$	$p = 0.943$	$r_s = 0.13$	$p = 0.838$	$r_s = 0.02$	$p = 0.915$	$r_s = 0.18$	$p = 0.563$
Talking											
Eyes vs mouth		$r_s = 0.05$	$p = 0.846$	$r_s = -0.14$	$p = 0.943$	$r_s = 0.10$	$p = 0.858$	$r_s = -0.05$	$p = 0.827$	$r_s = -0.01$	$p = 0.971$
People vs geometry											
Same-size											
People vs geometry		$r_s = -0.40$	$p < 0.05$	$r_s = -0.04$	$p = 0.943$	$r_s = -0.01$	$p = 0.963$	$r_s = 0.20$	$p = 0.643$	$r_s = 0.09$	$p = 0.749$
Embedded											
People vs geometry		$r_s = -0.41$	$p < 0.05$	$r_s = -0.09$	$p = 0.943$	$r_s = 0.25$	$p = 0.838$	$r_s = 0.09$	$p = 0.771$	$r_s = 0.19$	$p = 0.563$
Biological motion PLDs											
Upright vs inverted		$r = 0.17$	$p = 0.539$	$r = -0.09$	$p = 0.908$	$r = 0.15$	$p = 0.487$	$r = -0.28$	$p = 0.163$	$r = 0.03$	$p = 0.862$
Referential attention											
Target/person vs distractors		$r_s = -0.27$	$p = 0.200$	$r_s = -0.01$	$p = 0.956$	$r_s = 0.09$	$p = 0.858$	$r_s = 0.06$	$p = 0.785$	$r_s = 0.25$	$p = 0.485$
Target vs person/distractors		$r_s = 0.09$	$p = 0.653$	$r_s = 0.23$	$p = 0.943$	$r_s = -0.03$	$p = 0.881$	$r_s = 0.18$	$p = 0.643$	$r_s = 0.30$	$p = 0.485$
Target vs distractors		$r_s = -0.26$	$p = 0.205$	$r_s = 0.03$	$p = 0.943$	$r_s = 0.04$	$p = 0.881$	$r_s = 0.09$	$p = 0.771$	$r_s = 0.26$	$p = 0.485$

**Table 5.** Associations between gaze data to paired ROI (ratios) and scores on clinical/behavioural phenotyping measures for 51 infants with any available gaze tracking data. Significant values are in [bold]. All correlations are partial (controlling for within-stimulus % tracking rate) and  $p$  values are FDR corrected.  $r$  Pearson correlation,  $r_s$  Spearman's correlation, ROI Region of Interest, AOSI Autism Observation Scale for Infants, MSEL Mullen Scales of Early Learning, NVDQ Non-Verbal Developmental Quotient, VABS Vineland Adaptive Behavior Scales, ABC adaptive behaviour composite, AE age equivalence.

these same stimuli and, furthermore, no other association of assessed autism behaviours with gaze during human face, biological motion PLD, or referential attention trials.

The association of emerging autism behaviours with greater gaze towards geometry (alone, and vs. people) is consistent with a behavioural profile of social attention differences and circumscribed interests considered to be characteristic of autism<sup>1</sup> and the prior evidence of visual social attention differences in autism established on the basis of eye-tracking research using *Gazefinder*<sup>18</sup> and other technologies<sup>14</sup>. Indeed, heightened gaze to non-social ROIs has been shown to robustly discriminate autistic from non-autistic children and adults<sup>10,11</sup>, and a visual preference for geometric patterns in infancy identified as a promising biomarker for later autism outcome<sup>14</sup>.

The lack of observed association of autism behaviours and gaze towards eye vs. mouth ROIs within presentations of a human face is somewhat surprising given reduced eye contact is a cardinal feature of autism, and prior evidence of differential attention to facial features in autism<sup>10,11</sup>. However, research suggests that infants later diagnosed with autism initially show eye fixation patterns similar to typical levels, but this gradually declines by the age of 24 months, when reductions compared to children without autism become maximally apparent<sup>42</sup>. Similarly, the lack of association of autism behaviours with gaze towards the referential agent and target object (vs. distractors) in referential attention probes is surprising in light of the child and adolescent eye-tracking literature<sup>43–45</sup>. However, evidence of gaze-following differences in infants at higher likelihood or in infants who later received an autism diagnosis has not always been found<sup>46–48</sup>. Consequently, the *Gazefinder* human face and referential attention stimuli might prove more effective in identifying visual attention differences related to autism at later developmental stages<sup>17,18</sup>.

Moreover, while other behavioural studies suggest autistic people autism may be less sensitive to the social information conveyed in human movement (e.g., emotions, intentions)<sup>49</sup>, we found no evidence associating emerging autism behaviours with gaze to biological motion PLDs in this infant sample. This null finding is consistent with Fujioka et al.'s previous report following *Gazefinder* eye-tracking assessment with autistic and non-autistic adolescents/adults<sup>18</sup> but contrasts with other studies finding reduced fixation towards similar biological motion PLD stimuli in autistic children aged 2–7 years<sup>50–52</sup>. Limited utility of the specific biological motion PLD trials within the *Gazefinder* 'Scene 1S4' protocol may therefore be suggested for autism-related visual social attention assessment.

Finally, we observed virtually no association between gaze data and measures of cognitive/developmental functioning, suggesting the potential specificity of the *Gazefinder* people vs. geometry stimulus for the assessment of autism-related differences in visual attention. The moderate (albeit non-significant) association of assessed

receptive language with gaze towards a referential target aligns with the importance of shared and referential (i.e., dyadic and triadic) attention in language development<sup>20</sup>. The stronger correlation observed here, compared to the link between gaze at the same target and autism-related behaviours, indicates that this stimulus could be more useful in evaluating language abilities rather than autism behaviours, particularly during this early developmental stage.

### Limitations and future directions

This study adds to the growing body of evidence supporting the potential utility of eye-tracking technology in clinical autism assessment protocols, providing proof-of-concept for *Gazefinder* as a standardized, brief, user friendly—and highly scalable—option. Limitations of our study include the cross-sectional design and relatively small sample of participating infants. We have included repeated *Gazefinder* assessments within our prospective follow-up of this clinically-indicated cohort into early childhood, which will permit further evaluation of: (a) convergent validity against concurrent clinical/behavioural measures taken at later child ages (with larger participant samples, given the system was only available mid-way through the study recruitment phase and first assessment, reported here), and (b) predictive validity from earlier *Gazefinder* assessment to later clinical characterisation and diagnostic outcome.

Another limitation of note is that our implementation of *Gazefinder* was not within a true community clinical autism assessment service; rather, in university/research settings that mirrored a quality community practice. Further, we made minor but non-trivial adjustments to *Gazefinder* software and staff training during the course of the study (see Figures S1–S2). We have transparently reported on our experience of implementing *Gazefinder* within the broader protocol of clinical/behavioural assessments for our research program, and sought to closely mirror potential future clinical use of this system. Specifically, we deployed the *Gazefinder* GP-100EA apparatus off-the-shelf, implemented the standardised ~2-min ‘Scene 1S4’ autism assessment without modification, and attempted assessment with all consecutively enrolled infants from the point the system was available at both sites (i.e., regardless of medical history or clinical/behavioural presentation). Moreover, we analysed recorded gaze data to pre-specified ROIs, exported these data directly from the *Gazefinder* units without any post-processing, and retained all participants with any usable data (i.e., regardless of overall % tracking rate).

Finally, while the standardised stimulus presentation is a strength of the *Gazefinder* system, this also introduces the potential for undetectable presentation order effects which might plausibly explain some of our observed null findings. In particular, the diverging pattern of group mean-level preferential attention to eyes vs mouth of a static human face across initial and later presentations of this identical stimulus, and the lack of expected association of assessed autism behaviours and gaze towards less socially-salient ROIs across some trial types. As evident in the detailed ‘Scene 1S4’ stimulus parameters included here and in Supplementary Materials, some between-trial counterbalancing exists for left- vs. right-hand side of screen, even within the brief assessment protocol. Plausibly the costs of inadvertent presentation/order effects may be offset by the comparability benefits of employing a standardised animation across various study samples and populations.

The current data offer some support for the potential of *Gazefinder* with standardised ‘Scene 1S4’ stimulus sequence to offer a scalable objective assessment to support early autism detection. Employing sophisticated data-driven techniques, such as machine learning, could bring significant further insights from this technology and the data it derives automatically from infant and other viewers. Drawing on data from the most informative ROIs and gaze metrics—potentially including fixation durations but also saccade patterns, pupillary responses, etc.<sup>53,54</sup>—such approach could generate an autism prediction algorithm capable of assigning a probability or classification likelihood (e.g., high, moderate, low). Indeed, there has been a gradual increase in research adopting machine learning for eye-tracking data in attempt to classify autistic vs. non-autistic development, with a recent meta-analysis of 24 studies reporting pooled 81% classification accuracy<sup>55</sup>, underscoring the potential from integrating multiple eye-tracking indicators for autism diagnosis. The varied stimulus array and dataset offered by *Gazefinder* offers fertile ground for applying machine learning techniques and the current study represents an initial phase of this pursuit, laying the groundwork for further exploration and the potential scalable implementation of objective early autism identification.

### Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 14 September 2022; Accepted: 26 February 2024

Published online: 01 March 2024

### References

1. Association, A. P. *Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition (DSM-5)* 5th edn. (American Psychiatric Publishing, 2013).
2. Jeste, S. & Geschwind, D. Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat. Rev. Neurol.* **10**(2), 74–81 (2014).
3. Whitehouse, A. *et al. National Guideline for the Assessment and Diagnosis of Autism Spectrum Disorders in Australia* (Cooperative Research Centre for Living with Autism (Autism CRC), 2018).
4. Zwaigenbaum, L. *et al. Early identification of Autism Spectrum Disorder: Recommendations for practice and research. Pediatrics* **136**, s10–s40 (2015).
5. Dawson, G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Dev. Psychopathol.* **20**(3), 775–803 (2008).
6. Egger, H. *et al. Automatic emotion and attention analysis of young children at home: A ResearchKit autism feasibility study. NPJ Digit. Med.* **20**, 1 (2018).

7. Whitehouse, A. *et al.* *Interventions for Children on the Autism Spectrum: A Synthesis of Research Evidence* (Autism CRC, 2020).
8. Varcin, K. & Nelson, C. A developmental neuroscience approach to the search for biomarkers in autism spectrum disorder. *Curr. Opin. Neurol.* **29**(2), 123 (2016).
9. Zammarchi, G. & Conversano, C. Application of eye tracking technology in medicine: A bibliometric analysis. *Vision* **5**(4), 56 (2021).
10. Frazier, T. *et al.* A meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism. *J. Am. Acad. Child Adolesc. Psychiatry* **56**(7), 546–555 (2017).
11. Setien-Ramos, I. *et al.* Eye-tracking studies in adults with autism spectrum disorder: A systematic review and meta-analysis. *J. Autism Dev. Disord.* **20**, 25 (2022).
12. Ahmed, Z., & Jadhav, M. A review of early detection of autism based on eye-tracking and sensing technology. In *International Conference on Inventive Computation Technologies (ICICT)* (Institute of Electrical and Electronics Engineers: Coimbatore, India, 2020).
13. Mastergeorge, A., Kahathuduwa, C. & Blume, J. Eye-tracking in infants and young children at risk for autism spectrum disorder: A systematic review of visual stimuli in experimental paradigms. *J. Autism Dev. Disord.* **51**, 2578–2599 (2021).
14. Pierce, K. *et al.* Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity. *Biol. Psychiatry* **15**, 657–666 (2016).
15. Chawarska, K., Macari, S. & Shic, F. Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biol. Psychiatry* **74**(2), 195–203 (2013).
16. Jones, W. *et al.* Eye-tracking-based measurement of social visual engagement compared with expert clinical diagnosis of autism. *JAMA* **330**(9), 854–865 (2023).
17. Tsuchiya, K. *et al.* Diagnosing Autism Spectrum Disorder without expertise: A pilot study of 5- to 17-year-old individuals using Gazefinder. *Front. Neurol.* **11**, 603085 (2021).
18. Fujioka, T. *et al.* Gazefinder as a clinical supplementary tool for discriminating between autism spectrum disorder and typical development in male adolescents and adults. *Mol. Autism* **7**(19), 19 (2016).
19. Fukushima, S. *et al.* The feasibility of Gazefinder under 12 months of age infants. *Sci. Rep.* **11**, 10009 (2021).
20. Bottema-Beutel, K. Associations between joint attention and language in autism spectrum disorder and typical development: A systematic review and meta-regression analysis. *Autism Res.* **9**(10), 1021–1035 (2016).
21. Hudry, K. *et al.* Early language profiles in infants at high-risk for autism spectrum disorders. *J. Autism Dev. Disord.* **44**, 25 (2014).
22. Whitehouse, A. *et al.* Pre-emptive intervention versus treatment as usual for infants showing early behavioural risk signs of autism spectrum disorder: a single-blind, randomised controlled trial. *Lancet Child Adolesc. Health* **3**, 605–615 (2019).
23. Whitehouse, A. *et al.* Effect of preemptive intervention on developmental outcomes among infants showing early signs of autism: A randomized clinical trial of outcomes to diagnosis. *JAMA Pediatr.* **175**(11), e213298 (2021).
24. Barbaro, J. *et al.* Diagnostic accuracy of the Social Attention and Communication Surveillance-Revised with Preschool tool for early autism detection in very young children. *JAMA Netw. Open* **5**, 25 (2022).
25. Bryson, S. *et al.* The Autism Observation Scale for Infants: Scale development and reliability data. *J. Autism Dev. Disord.* **38**, 731–738 (2008).
26. Hudry, K. *et al.* Performance of the Autism Observation Scale for Infants with community-ascertained infants showing early signs of autism. *Autism* **25**(2), 490–501 (2021).
27. Gammer, I. *et al.* Behavioural markers for autism in infancy: Scores on the Autism Observational Scale for Infants in a prospective study of at-risk siblings. *Infant Behav. Dev.* **38**, 107–115 (2015).
28. Zwaigenbaum, L. *et al.* Behavioral manifestations of autism in the first year of life. *Int. J. Dev. Neurosci.* **23**, 143–152 (2005).
29. Whitehouse, A. *et al.* A randomised controlled trial of an iPad-based application to complement early behavioural intervention in Autism Spectrum Disorder. *J. Child Psychol. Psychiatry* **59**(9), 1042–1052 (2017).
30. Mullen, E. *Mullen Scales of Early Learning* (American Guidance Service, 1995).
31. Sparrow, S., Cicchetti, D. & Balla, D. *The Vineland Adaptive Behavior Scales, 2nd Edition (VABS-II)* (NCS Pearson Inc, 2005).
32. Gartstein, M. & Rothbart, M. Studying infant temperament via the revised Infant Behavior Questionnaire. *Infant Behav. Dev.* **26**(1), 64–86 (2003).
33. Chetcuti, L. *et al.* Subgroups of temperament associated with social-emotional difficulties in infants with early signs of autism. *Autism Res.* **20**, 20 (2020).
34. Corporation, J. *Gazefinder Instruction Manual: Gaze Tracking and Diagnostic System Model GP-100EA* (Corporation, J, 2019).
35. Corporation, J. *Gazefinder Instruction Manual: Operation 68* (Corporation, J, 2022).
36. Fujioka, T. *et al.* Developmental changes in attention to social information from childhood to adolescence in autism spectrum disorders: a comparative study. *Mol. Autism* **11**, 25 (2020).
37. Baio, J. *et al.* Prevalence of Autism Spectrum Disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveill. Summaries* **67**, 1–23 (2018).
38. Bent, C., Barbaro, J. & Dissanayake, C. Mapping the diagnosis of Autism Spectrum Disorders in children under 7 years in Australia: 2010–2012. *Med. J. Aust.* **202**, 317–320 (2015).
39. Kleinman, J. *et al.* Diagnostic stability in very young children with Autism Spectrum Disorders. *J. Autism Dev. Disord.* **38**, 606–615 (2008).
40. Elsabbagh, M. *et al.* Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Curr. Biol.* **22**, 338–342 (2012).
41. Kato, M. & Konishi, Y. Where and how infants look: The development of scan paths and fixations in face perception. *Infant. Behav. Dev.* **36**(1), 32–41 (2012).
42. Jones, W. & Klin, A. Attention to eyes is present but in decline in 2–6 month-olds later diagnosed with autism. *Nature* **504**, 427–431 (2013).
43. Billeci, L. *et al.* An integrated EEG and eye-tracking approach for the study of responding and initiating joint attention in Autism Spectrum Disorders. *Sci. Rep.* **7**, 1 (2017).
44. Caruana, N. *et al.* Joint attention difficulties in autistic adults: An interactive eye-tracking study. *Autism* **22**(4), 502–512 (2018).
45. Zhang, J. *et al.* RJAfinder: An automated tool for quantification of responding to joint attention behaviors in autism spectrum disorder using eye tracking data. *Front. Neurosci.* **16**, 25 (2022).
46. Bedford, R. *et al.* Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *J. Autism Dev. Disord.* **42**, 2208–2218 (2012).
47. Billeci, L. *et al.* Disentangling the initiation from the response in joint attention: An eye-tracking study in toddlers with autism spectrum disorders. *Transl. Psychiatry* **6**(5), e808 (2016).
48. Nystrom, P. *et al.* Joint attention in infancy and the emergence of autism. *Biol. Psychiatry* **86**(8), 631–638 (2019).
49. Kaiser, M. & Shiffrar, M. The visual perception of motion by observers with autism spectrum disorders: A review and synthesis. *Psychon. Bull. Rev.* **16**, 761–777 (2009).
50. Annaz, D. *et al.* Young children with autism spectrum disorder do not preferentially attend to biological motion. *J. Autism Dev. Disord.* **42**(3), 401–408 (2012).
51. Klin, A. *et al.* Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* **459**, 257–261 (2009).

52. Wang, L. *et al.* Children with autism spectrum disorders are less proficient in action identification and lacking a preference for upright point-light biological motion displays. *Res. Autism Spectrum Disord.* **11**, 63–76 (2015).
53. Mahanama, B. *et al.* Eye movement and pupil measures: A review. *Front. Comput. Sci.* **3**, 733531 (2022).
54. Sasson, N. & Elison, J. Eye tracking young children with autism. *J. Visual. Exp.* **61**, e3675 (2012).
55. Wei, Q. *et al.* Machine learning based on eye-tracking data to identify Autism Spectrum Disorder: A systematic review and meta-analysis. *J. Biomed. Inform.* **20**, 25 (2022).

## Acknowledgements

We thank the infants and caregivers for their time commitment to participate in this study and Sarah Fritsche, Natalie Mizzi, Ashley Rattenbury, Megan Harrap, and Sarah Pillar for their research support. We acknowledge the prior development work for *Gazefinder* and the stimuli used in this research by Prof. Kenji J. Tsuchiya (Hama-matsu University School of Medicine, Japan) and Prof. Taiichi Katayama (Osaka University, Japan), and advice and support from Taeko Ebihara, Masaru Ninomiya and Atsushi Saito (JVCKENWOOD Corporation, Japan).

## Author contributions

L.C. contributed to the acquisition, analysis, and interpretation of data, and to drafting the article. K.V. contributed to the acquisition and interpretation of data, and to critically revising the article for important intellectual content. M.B. contributed to the acquisition of data, and to critically revising the article for important intellectual content. J.S. contributed to critically revising the article for important intellectual content. C.B. contributed to critically revising the article for important intellectual content. A.J.O.W. contributed to conception and design, interpretation of data, and critically revising the article for important intellectual content. K.H. contributed to conception and design, to analysis and interpretation of data, and to drafting the article. All authors approved the final version to be submitted/published.

## Funding

This research was supported with funding by JVCKENWOOD Corporation, Japan, to KH and AJOW. AJOW is supported by an Investigator Grant (1173896) from the National Health and Medical Research Council, Australia. The larger study alongside which this project was aligned was funded by the La Trobe University Understanding Disease Research Focus Area, the Cooperative Research Centre for Living with Autism (Autism CRC) established and supported by the Australian government's Cooperative Research Centre program, the Western Australian Children's Research Fund, and the Angela Wright Bennett Foundation. The funder had no role in the design and conduct of the study, manuscript drafting, or decision to submit the manuscript for publication.

## Competing interests

KH and AJOW received research funding from JVCKENWOOD Corporation, including for the conduct of this study. The other authors have no conflicts of interest to disclose.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-55643-z>.

**Correspondence** and requests for materials should be addressed to K.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024