



REPORT

Altered face scanning and impaired recognition of biological motion in a 15-month-old infant with autism

Ami Klin¹ and Warren Jones^{1,2}

1. Yale Child Study Center, Yale University School of Medicine, USA

2. Interdepartmental Neuroscience Program, Yale University School of Medicine, USA

Abstract

Mounting clinical evidence suggests that abnormalities of social engagement in children with autism are present even during infancy. However, direct experimental documentation of these abnormalities is still limited. In this case report of a 15-month-old infant with autism, we measured visual fixation patterns to both naturalistic and ambiguous social stimuli: video scenes of a caregiver and point-light animations of human action. Results suggested that viewing patterns of the child with autism were driven by the physical contingencies of the stimuli rather than by their social context. If corroborated in larger studies, this observation would advance the hypothesis that mechanisms of social development which rely on preferential engagement with socially contingent conspecifics – and which emerge in the very first weeks of life in typical infants – are developmentally derailed in children with autism.

Introduction

In this case report, we describe a serendipitous observation made during experiments examining mental representations of human action in a well-characterized 15-month-old infant with autism (Klin, Chawarska, Paul, Rubin, Morgan, Wiesner & Volkmar, 2004). Autism is a congenital and lifelong disorder of socialization, disrupting the development of interpersonal engagement, communication, and aspects of cognition (Volkmar, Lord, Bailey, Schultz & Klin, 2004). One theoretical approach postulates that the origin of these abnormalities lies in the disruption of innate mechanisms predisposing infants to seek social stimuli (Dawson, Meltzoff, Osterling, Rinaldi & Brown, 1998; Klin, 1991; Klin, Jones, Schultz & Volkmar, 2003; Schultz, 2005). To date, however, little experimental evidence is available on infants with autism because these children are typically identified only at later stages of childhood (Volkmar, Chawarska & Klin, 2005). Nevertheless, mounting clinical evidence revealed in systematic analyses of home movies filmed before the child was diagnosed (Osterling, Dawson & Munson, 2002; Werner, Dawson, Osterling & Dinno, 2000), as well as in retrospective and prospective studies (Volkmar, Stier & Cohen, 1985 and Zwaigenbaum, Bryson, Rogers, Roberts, Brian & Szatmari, 2005, respectively), suggest that abnormalities are present in the first year of life, prior to the time in which developmental concerns typically lead to clinical referrals (Chawarska, Klin, Paul, Hannigen, Dichtel & Volkmar, 2007). Results of

the experiments reported here support this view and advance the hypothesis that infants with autism may not view a face with the special status ascribed to persons or conspecifics, an ability found in typical babies at the age of 4 days (Farroni, Csibra, Simion & Johnson, 2002).

Comprehensive assessment and characterization information for the child in this report is given in an earlier publication (Klin *et al.*, 2004). This child was first seen at 15 months, and then re-evaluated at 23 and 34 months for confirmatory diagnosis and characterization. At the chronological age of 15 months (at the time of the experiments reported here), her nonverbal skills, fine and gross motor skills were at age level, whereas language skills were delayed at approximately the 9-month level. She met criteria for autism on both standardized diagnostic instruments and clinician-assigned diagnosis. Genetic and neurological evaluations were essentially non-contributory, although this child had an older sibling with autism.

Procedure and results

We presented the infant with autism with eight point-light animations (Johansson, 1973; Allison, Puce & McCarthy, 2000), each emulating social experiences in infancy and lasting an average of 30 seconds. The capacity for recognizing point-light displays of this nature as exemplars of 'biological motion' (e.g. distinguishing

Address for correspondence: Ami Klin, Yale Child Study Center, 230 South Frontage Road, New Haven, CT 06520, USA; e-mail: ami.klin@yale.edu

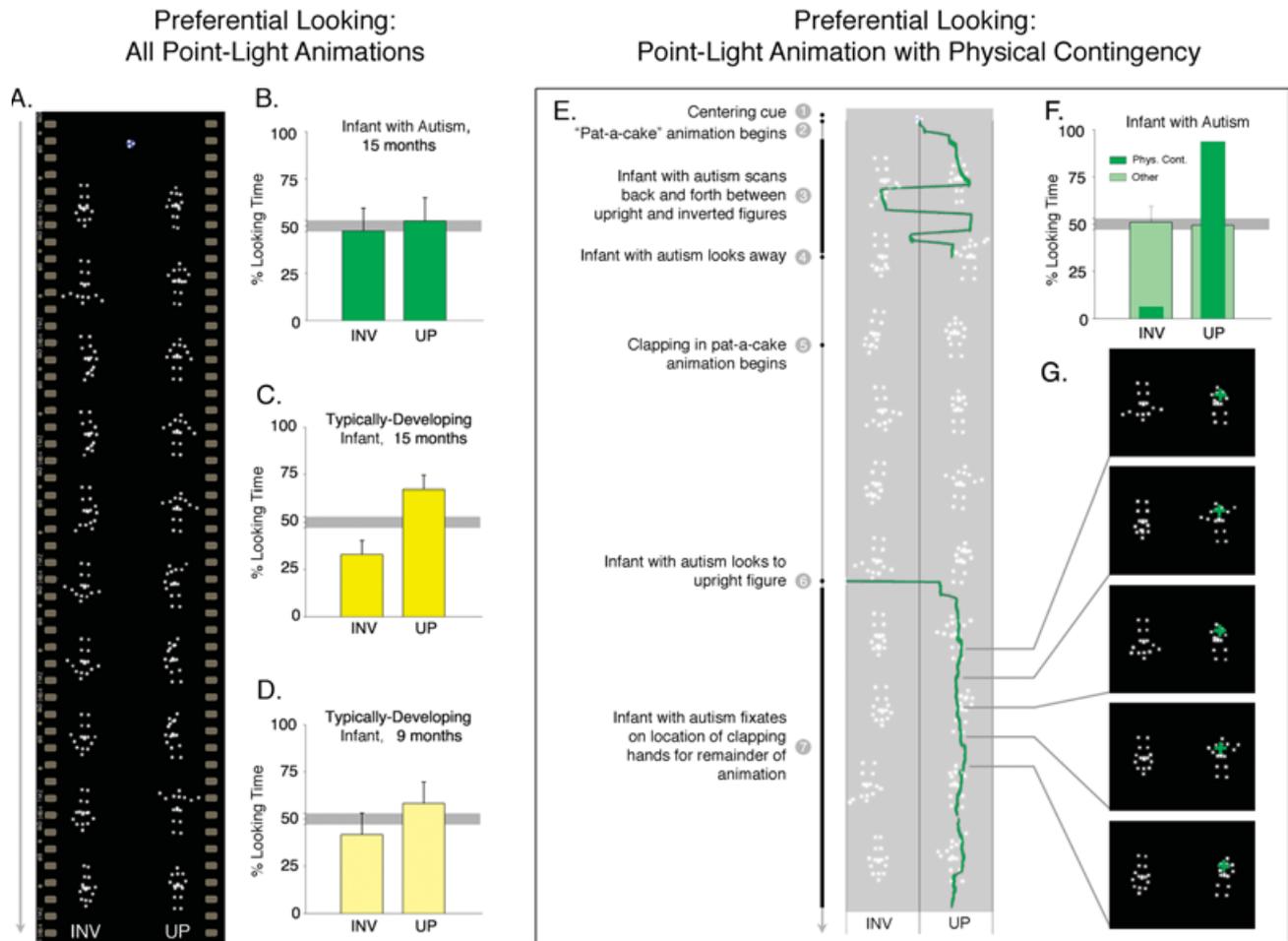


Figure 1 Preferential looking to point-light social animations. *A.* Example stimuli. Each animation showed one upright (UP) and one inverted (INV) point-light figure with an accompanying audio soundtrack (which matched the actions of the upright figure). The upright figure acted out childhood games (e.g. playing peek-a-boo, singing a nursery rhyme). The upright and inverted figures were identical except that the inverted figure was rotated 180 degrees and played in reverse order. Movie presentation was randomized and counterbalanced so that the upright figure appeared on the left side of the screen as often as on the right. *B.* Chance-level fixation time percentages on upright vs. inverted figures for 15-month-old infant with autism. *C.* Preferential fixation on upright figures for nonverbal mental age-matched, typically developing control. *D.* Preferential fixation on upright figures for verbal mental age-matched, typically developing control. Bar graphs are percentage of total fixation time, with SEM error bars. Horizontal gray zone denotes percentages which are not significantly different from chance. *E.* One point-light animation ('pat-a-cake') presented a physical contingency: when the dots marking the hands came together, the sound of a clap could be heard. The horizontal fixation position of the infant with autism is plotted over time in green (overlaid on example stills from the pat-a-cake animation). Before the physical contingency, scanning of the infant with autism is random; after onset of the physical contingency, the infant with autism fixates exclusively on the upright figure. *F.* Percentage of fixation on upright vs. inverted figure for pat-a-cake animation (with physical contingency) vs. all other animations (without physical contingencies). *G.* Screenshots of data from infant with autism showing fixation on the region where clapping occurs.

moving dots depicting a walking person from dots moving randomly) has been shown in 3-month-old infants (Fox & McDaniel, 1982). The animations appeared on a computer screen with an upright point-light figure on one half of the screen and an inverted (or upside-down) figure on the other (with counterbalanced presentation to pre-empt side bias) (Figure 1A) (Pavlova & Sokolov, 2000). The animations were accompanied by the vocal audio track of the social event (e.g. saying hello and waving to get the baby's attention, playing 'peek-a-boo', etc.). Evidence for mental representation of a given social action was measured as a function of the child's

preferential viewing patterns: preferential looking towards the upright figure showed evidence of the matching of sound effects with a mental template elicited by the animation (Klin *et al.*, 2003).

The child's pattern of preferential looking was measured with eye-tracking equipment, with data obtained at the rate of 60 samples per second. A typically developing 15-month-old infant matched on nonverbal mental age (NVMA) and a typically developing 9-month-old baby matched on verbal age (VMA) completed the same procedure for comparison. Nonverbal and verbal mental ages were obtained with the *Mullen Scales of Early*

Learning (Mullen, 1995). To analyze preferential viewing, non-fixation data (saccades, blinks, and off-screen fixations) were identified from within the 3 minutes and 57 seconds of total viewing data (saccades, blinks, and off-screen visual fixations were identified by measurements of eye velocity, eyelid closure, and fixation location coordinates beyond the bounds of the video screen as detailed in Klin, Jones, Schultz, Volkmar & Cohen, 2002). Saccades, blinks, and off-screen fixations comprised 88.3 seconds of data for the infant with autism, and 31.2 and 62.4 seconds, respectively, for the NVMA- and VMA-matched children. The children did not differ significantly in their time spent blinking, but the child with autism did have a significantly higher rate of saccading and a significantly greater time spent looking off-screen relative to the typically developing controls. In all children, lost data comprised less than one-quarter of the total viewing time, which for studies of infant viewing represents a considerable improvement (e.g. Hunnius & Geuze, 2004). In the child with autism, both the increase in saccades (looking back and forth between the two animations repeatedly) and the increase in offscreen fixations (looking to other areas in the room) is in agreement with the hypothesis that the child with autism had less comprehension of the animations and, perhaps consequently, less interest in them.

Preferential viewing in our design was a binary choice (upright vs. inverted). To determine viewing preferences significantly different from chance, we modeled total viewing time as a binomial distribution. Average possible viewing time, in number of frames of video fixated by the infants (a figure akin to the number of attempted flips in a coin toss), was 5293. Modeling the binary outcome for this number of trials indicates that results between 47% and 53% should be considered random viewing (at $p < .01$). Summaries for the eight animations indicated random choice by the infant with autism: 52.8% and 47.2% for upright and inverted, respectively. The NVMA and the VMA children, however, both showed strong visual preference for upright relative to inverted animations: 66.1 vs. 33.9% for the NVMA child and 58.8% vs. 41.2% for the VMA child (Figure 1B, 1C, 1D; also, see supplementary video S1 available online).

Preferential viewing for the infant with autism was random when summed across all animations. However, one animation yielded outlier results: in that animation the infant with autism spent 93.9% of fixation time on the upright figure vs. only 6.1% on the inverted (Figure 1F). Removal of this outlier actually made the child's overall results 'more' random: 48.6% on upright vs. 51.4% on inverted. Closer inspection of the outlier animation indicated an interesting experimental confound. In that animation, in which an adult acts out a game of 'pat-a-cake', there was a causal physical contingency between image and sound: two of the animation's points of light (the figure's 'hands') collide on-screen as the sound of a clap is heard. Of all the animations presented, the pat-a-cake animation is the only one in which

the movement of point-lights directly coincides with the physical production of a sound. In the other animations, only speech sounds are present. While the speech sounds are socially contingent (that is, the speech sounds fit the animation by way of a common social context), the speech sounds themselves are not the physical consequence of the point-lights' movement. During pat-a-cake, this physical contingency only exists in the upright animated figure because the action of the inverted figure is playing in reverse (so the inverted figure's hand clapping does not make a sound). This reversal (as can be seen in Figure 1A) was done to prevent the two figures from mirroring one another.

During the pat-a-cake animation, the looking pattern for the infant with autism indicates that she was sensitive to the physical synchrony of the clapping. Figure 1E shows this child's visual scanning and fixation unfolding over time. Her sensitivity to the audiovisual synchrony is highlighted both by her shift in fixation after the onset of clapping (looking back to the screen) and by her sustained attention to the exact location where the clapping occurs for the remainder of the animation (1E and 1G; also, see supplementary video S2 available online). This viewing pattern suggests a relatively advanced capacity for cross-modal matching of audio and visual signals, and this ability – to match contingencies of a physical nature – contrasts with her inability to match contingencies in the other animations which depended instead upon social context rather than physical causation.

Following these observations, we questioned whether this child's sensitivities to physically contingent vs. socially contingent stimuli might also result in atypical visual fixation patterns when the child with autism looked at people. Typical infants prefer to focus on the eye region of the face from at least the age of 3 months (Haith, Bergman & Moore, 1977), consistent with the idea that the eyes convey the most social-affective information (Emery, 2000). In contrast, this child's sensitivity to information which was physically synchronous in the point-light animations, together with the cross-modal physical contingency of speech sounds and lip movements, predicted that when looking at videos of a caregiver the infant with autism would preferentially fixate on the mouth region of the face.

To explore this, the three children were presented with 10 video scenes (full-screen video plus audio track), each lasting approximately 20 seconds and showing a female actor playing the role of caregiver: looking directly into the camera and entreating the viewing infant by engaging in childhood games (e.g. playing peek-a-boo). Visual fixation patterns were measured with eye-tracking equipment as in the first experiment (Figure 2A–C; also, see supplementary video S3 available online). On-screen fixations were coded relative to four regions of interest: eyes, mouth, body (neck and contours around eyes and mouth such as hair), and object (surrounding inanimate stimuli) (Figure 2E). To analyze fixation patterns, saccades, blinks, and off-screen fixations were identified from

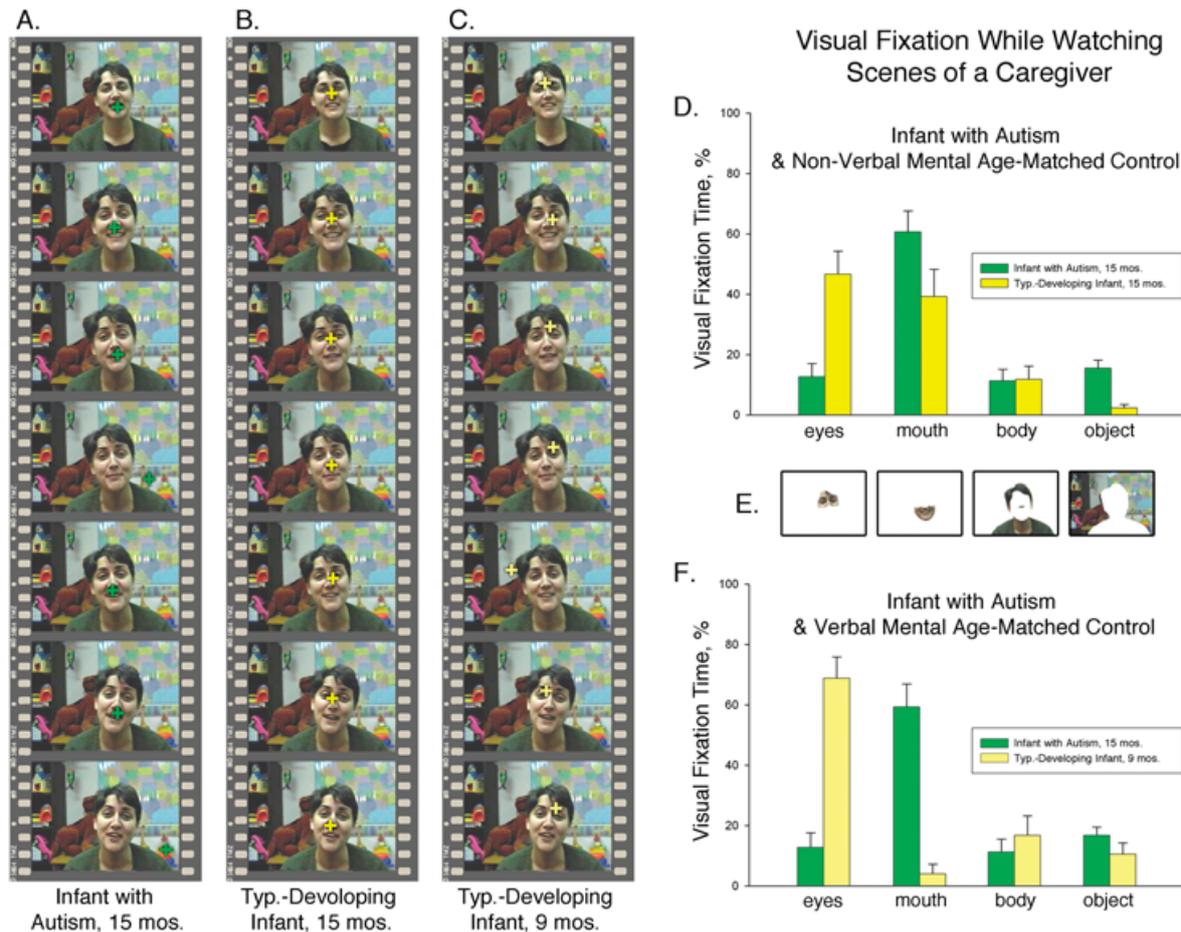


Figure 2 Visual fixation while watching scenes of a caregiver. A. Screenshots of data from infant with autism (age 15 months). B. Screenshots of data from typically developing control matched on nonverbal mental age (age 15 months). C. Screenshots of data from typically developing control matched on verbal mental age (age 9 months). D. Percentage of visual fixation time on eyes, mouth, body, and object regions for infant with autism and nonverbal mental age-matched control. E. Example coding regions for one frame of video. F. Percentage of visual fixation time on eyes, mouth, body, and object regions for infant with autism and verbal mental age-matched control. Bar graphs are percentage of total fixation time, with SEM error bars.

within the 3 minutes and 26 seconds of total viewing data. Saccades, blinks, and off-screen fixations accounted for 59.6 seconds of data for the infant with autism, and 37.6 and 88.2 seconds for the NVMA- and VMA-matched children, respectively. The remaining fixation data were analyzed for looking time at the four regions of interest. As predicted, the infant with autism showed viewing preference for the mouth region rather than the eye region of the face relative to the NVMA infant (Figure 2D). This relative preference was even stronger in comparison with the VMA-matched child (Figure 2F).

Discussion

Our conclusions from this case study are necessarily tempered because they are based on a single case relative to comparison children. Substantiation of the hypotheses awaits replication in larger studies using *a priori* experi-

mental manipulations to replicate the post-hoc findings here. This case presents intriguing insights, however, and – by way of an experimental confound and a serendipitous observation – opens an avenue of interesting follow-up studies.

The fixation patterns of the child with autism could be explained by a variety of factors. Some fixation on the mouth region (although still less than on the eye region) is expected in infants learning to speak given the bimodal (auditory and visual) perception of speech acquisition (Kuhl & Meltzoff, 1982). Given the state of language development in the infant with autism, however, we expect that this would play only a minimal role in her fixation patterns. We would expect this to be a stronger factor in the older, NVMA-matched infant, and we would expect that the VMA-matched infant would provide an appropriate comparison point for fixation patterns which are less influenced by language development. The marked differences in fixation patterns between the child with autism and both the VMA-matched

child and the NVMA-matched child suggest a different mechanism.

The fixation patterns for the infant with autism could also be explained as a function of preferential attention to an area of maximal localized movement (the mouth in motion). However, given the results of the first experiment, where this infant showed sensitivity to a cross-modal (audiovisual) physical event, it seems more likely that if motion is an attraction, synchronous sound and motion are still more attractive targets for the child's attention. Future experimentation should disambiguate this finding. In either case, these results suggest that the infant with autism attended to physical contingencies rather than social contingencies: looking more at areas in the environment which exhibit physical causality while looking less at areas which are typically of maximal social value (such as the expressive and entreating eye gaze of a caregiver).

Preferential looking to the mouth region of the face relative to the eye region has also been shown in older and higher-functioning (i.e. normative IQ) individuals with autism (Klin *et al.*, 2002). Although the results of this case study show similar patterns of looking, it seems unlikely that the underlying mechanism at these different developmental time points would be the same. In older individuals, correlation between mouth fixation and social competence suggested that looking at the mouth was a compensatory strategy (higher fixation on mouths correlated with higher levels of social ability and lower level of autistic symptomatology as measured through standardized instruments; Klin *et al.*, 2002). We hypothesized that these individuals focused on the mouth in an attempt to rely on language as a relatively concrete inroad into understanding naturalistic social situations (given their difficulty processing intentional and affective cues conveyed through the eye region; Baron-Cohen, Campbell, Karmiloff-Smith, Grant & Walker, 1995). In that study, all participants had good language skills alongside hallmark deficits in processing of nonverbal communication (such as processing of prosody, facial expressions, body language and gestures).

The child in the present case report, however, is very different: her language skills were at the 9-month level (Klin *et al.*, 2004). It seems unlikely, therefore, that this infant with autism would focus on mouths for the same reason as older individuals with autism or that she would receive the same compensatory information as language-proficient individuals. Instead, this may be evidence of a much earlier time point in the developmental trajectory of how children with autism learn about the world. These children may focus first on the mouth because of its physically contingent properties, and then over time learn about language by parsing the physical relationship between motion and sound. This theoretical speculation suggests an alternative learning path for language acquisition, the results of which would dovetail both with language delays in individuals with autism and with persistent language difficulties (such as with

prosody or with the understanding of sarcasm or non-literal speech – instances which are more dependent upon language as part of a social context than language as a literal mapping of sound to singular meaning).

Findings in this case report contrast with another case study reported in the literature (Baron-Cohen, Scott, Wheelwright, Johnson, Bisarya, Desai & Ahluwalia, 2006). That case study employed a different methodology and the two infants had very different clinical presentations. In our study, the infant exhibited prototypical autism (Klin *et al.*, 2004), whereas in the Baron-Cohen and colleagues' study, the child's condition was described with the term Asperger Syndrome, which likely represented a milder form of autism. The different results could reflect different developmental pathways leading to varied manifestations of autism spectrum disorders. Further consideration of this question will benefit from larger studies involving a fuller range of autism manifestations and more control participants.

Efforts are under way to employ the methods used here with a larger sample of equally young children with autism. If corroborated in larger studies, these results would suggest that infants with autism are more sensitive to physical contingencies in their environment than to social-affective contingencies. In dyadic social engagement, this would be manifest through decreased fixation on the eye region of the face and increased fixation on areas with high physical synchrony. In the most extreme interpretation of this hypothesis, this pattern of looking would suggest seeing the world, and even people, as a collection of physical contingencies, unmoored from their social context or adaptive relevance. Living in such a world would likely have a profound impact on the development of the social mind and on the organization of the social brain (Johnson, 2001).

Acknowledgements

This work was supported by a grant from the National Institutes of Mental Health (U54-MH66494), as well as by Autism Speaks and the Simons Foundation. We also wish to thank Katelin Carr, Phillip Gorrindo, and Amanda Blank for their help in this research project.

References

- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.
- Baron-Cohen, A., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13*, 379–398.
- Baron-Cohen, S., Scott, F., Wheelwright, S., Johnson, M., Bisarya, D., Desai, A., & Ahluwalia, J. (2006). Can Asperger syndrome be diagnosed at 26 months old? A genetic high-risk single-case study. *Journal of Child Neurology*, *21*, 351–356.

- Chawarska, K., Klin, A., Paul, R., Hannigen, S., Dichtel, E., & Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **37** (1), 62–72.
- Dawson, G., Meltzoff, A., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, **28**, 479–485.
- Emery, N.J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Review*, **24**, 581–604.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M.H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, **99** (14), 9602–9605.
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, **218**, 486–487.
- Haith, M.M., Bergman, T., & Moore, M. (1977). Eye contact and face scanning in early infancy. *Science*, **218**, 179–181.
- Hunnius, S., & Geuze, R.H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: a longitudinal study. *Infancy*, **6**, 231–255.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, **14**, 201–211.
- Johnson, M. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, **2**, 475–483.
- Klin, A. (1991). Young autistic children's listening preferences in regard to speech: a possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, **21**, 29–42.
- Klin, A., Chawarska, K., Paul, R., Rubin, E., Morgan, T., Wiesner, L., & Volkmar, F.R. (2004). Autism in a 15-month-old child. *American Journal of Psychiatry*, **161** (11), 1981–1988.
- Klin, A., Jones, W., Schultz, R.T., Volkmar, F.R., & Cohen, D.J. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, **59** (9), 809–816.
- Klin, A., Jones, W., Schultz, R.T., & Volkmar, F.R. (2003). The enactive mind – from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society London, B, Biological Sciences*, **358**, 345–360.
- Kuhl, P.K., & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science*, **218**, 1138–1141.
- Mullen, E. (1995). *Mullen Scales of Early Learning*. AGS Edition. Circle Pines, MN: American Guidance Service, Inc.
- Osterling, J., Dawson, G., & Munson, J. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Developmental Psychopathology*, **14**, 239–251.
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception and Psychophysics*, **62**, 889–899.
- Schultz, R.T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience*, **23**, 125–141.
- Volkmar, F.R., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology*, **56**, 315–36.
- Volkmar, F.R., Lord, C., Bailey, A., Schultz, R.T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, **45**, 1–36.
- Volkmar, F.R., Stier, D.M., & Cohen, D.J. (1985). Age of recognition of Pervasive Developmental Disorder. *American Journal of Psychiatry*, **142**, 1450–1452.
- Werner, E., Dawson, G., Osterling, J., & Dinno, H. (2000). Recognition of autism spectrum disorder before one year of age: a retrospective study based on home videotapes. *Journal of Autism and Developmental Disorders*, **30**, 157–162.
- 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*, **23**, 143–152.

Received: 28 August 2006

Accepted: 26 February 2007

Supplementary Material

This set of supplementary materials is comprised of three QuickTime videos (audio and visual) providing a sample of video data described in the manuscript.

Video Clip S1 complements Figure 1B, 1C, and 1D. It shows:

- Top left quadrant: sample of point-light animations presented to children;
- Top right quadrant: sample of visual fixation patterns of 15-month-old infant with autism;
- Bottom left quadrant: sample of visual fixation patterns of 15-month-old typically developing infant (matched to the infant with autism on nonverbal mental age);
- Bottom right quadrant: sample of visual fixation patterns of 9-month-old typically developing infant (matched to the infant with autism on verbal mental age).

Note: In these illustrations, Upright point-light animations are in red and Inverted point-light animations are in green. Stimuli were presented to children as white point lights against the black background as can be seen in the top left quadrant.

Video Clip S2 complements Figure 1E and 1G. It shows:

- Top left quadrant: sample of pat-a-cake animation presented to children;
- Top right quadrant: sample of visual fixation patterns of 15-month-old infant with autism watching the pat-a-cake animation;
- Bottom left quadrant: sample of visual fixation patterns of 15-month-old typically developing infant watching the pat-a-cake animation;
- Bottom right quadrant: sample of visual fixation patterns of 9-month-old typically developing infant watching the pat-a-cake animation.

Note: In these illustrations, Upright point-light animations are in red and Inverted point-light animations are in green. Stimuli were presented to children as white point lights against the black background as can be seen in the top left quadrant.

Video Clip S3 complements Figure 2A, 2B, and 2C. It shows:

- *Top left quadrant: sample of caregiver approaches;*
- *Top right quadrant: sample of visual fixation patterns of 15-month-old infant with autism watching caregiver approaches;*
- *Bottom left quadrant: sample of visual fixation patterns of 15-month-old typically developing infant watching caregiver approaches;*
- *Bottom right quadrant: sample of visual fixation patterns of 9-month-old typically developing infant watching caregiver approaches.*

Note: The cross on the screen indicates coordinates of visual fixation. The color of the cross changes to reflect the region of interest (eyes, mouth, body, object). It also changes to

white during saccades, black during blinks, and there is a gray square in the upper left corner (in each quadrant) when the child's fixation is off screen and data are lost.

This material is available as part of the online article from:
<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1467-7687.2007.00608.x>

(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.