

# **Reduced Orienting to Audiovisual Synchrony in Infancy Predicts Autism Diagnosis at Three Years of Age**

Terje Falck-Ytter<sup>1,2,3\*</sup>, Pär Nyström<sup>1</sup>, Gustaf Gredebäck<sup>1</sup>, Teodora Gliga<sup>4</sup>, Sven Bölte<sup>2,3</sup> and  
the EASE team.

<sup>1</sup>Department of Psychology, Uppsala University, Uppsala, Sweden

<sup>2</sup>Center of Neurodevelopmental Disorders at Karolinska Institutet (KIND), Karolinska Institutet, Stockholm, Sweden

<sup>3</sup>Child and Adolescent Psychiatry Stockholm, Center for Psychiatry Research, Stockholm County Council, Sweden

<sup>4</sup>Centre for Brain and Cognitive Development, Birkbeck, University of London, UK

\*corresponding author

SHORT TITLE: Visual preferences in infants with later autism

**Word count:** 6500

**Figures:** n=3; **Tables:** n=2; **Supplementary material** n=0.

*Abstract:*

**Background:** Effective multisensory processing develops in infancy and is thought to be important for the perception of unified and multimodal objects and events. Previous research suggests impaired multisensory processing in autism, but its role in the early development of the disorder is yet uncertain. Here, using a prospective longitudinal design, we tested whether reduced visual attention to audiovisual synchrony is an infant marker of later-emerging autism diagnosis.

**Methods:** We studied 10-month-old siblings of children with autism using an eye tracking task previously used in studies of preschoolers. The task assessed the effect of manipulations of audiovisual synchrony on viewing patterns while the infants were observing point light displays of biological motion. We analyzed the gaze data recorded in infancy according to diagnostic status at three years of age (DSM-5).

**Results:** Ten-month-old infants who later received an autism diagnosis did not orient to audiovisual synchrony expressed within biological motion. In contrast, both infants at low risk and high-risk siblings without autism at follow up had a strong preference for this type of information. No group differences were observed in terms of orienting to upright biological motion.

**Conclusions:** This study suggests that reduced orienting to audiovisual synchrony within biological motion is an early sign of autism. The findings support the view that poor multisensory processing could be an important antecedent marker of this neurodevelopmental condition.

**Keywords:** autism spectrum disorder; infancy; multisensory processing; biological motion; biomarker; scientific replication

Autism Spectrum Disorder (in this report referred to as autism for simplicity) is a common and impairing neurodevelopmental condition. Currently, no effective treatment for core symptoms is in reach, but it is hoped that research into the early signs of the disorder may generate new leads to malleable behavioral and brain processes associated with the later emergence of clinical symptoms. Efficiently linking information from different sensory channels is critical for the perception of a unified, multisensory world. Multisensory information is prioritized from an early age in humans as well as other species (Bahrick, Liekliter, & Flom, 2004; Stein & Meredith, 1993), and it has been proposed that problems with multisensory processing could be a core feature in the early development of autism (Bahrick, 2010). Although this hypothesis has so far not been tested, data from children and adults with autism as well as animal models have generally found that multisensory processing is impaired in autism (Brandwein et al., 2012; Gogolla, Takesian, Feng, Fagiolini, & Hensch, 2014; Russo et al., 2010; Stevenson et al., 2014). For example, comparing children with autism to age-, IQ- and gender-matched controls, Brandwein and colleagues (2012) found striking reductions in audiovisual facilitation at the behavioral level, results that were paralleled by diminished amplitude and less wide-spread responses to audiovisual information in electroencephalography recordings from the same experimental sessions. Gogolla et al. (2014) found that that despite different etiology at a genetic and molecular level, several autism mouse models share a reduced capacity to link information from different senses. These disruptions reflected GABA-mediated excitation-inhibition imbalance, and could be rescued through early pharmacological intervention. Similar disruptions in GABA-circuits have been linked to activation in the visual cortex during auditory stimulation, indicative of suboptimal specialization of sensory networks (Hattori et al., 2017).

The few studies that exist on multisensory processing in very young children with autism are inconclusive, particularly when it comes to sensitivity to audiovisual synchrony (*i.e.*, when changes in audio volume coincide with changes in visual velocity). One study (Klin, Lin, Gorrindo, Ramsay, & Jones, 2009) found that compared to typically developing controls, two-year-olds with autism spent *more* time looking at audiovisual synchronous events, expressed within biological motion (*i.e.*, social information that is prioritized by typically developing infants and children; Simion, Regolin, & Bulf, 2008). This suggested that, in this context, young children with autism are not only able to efficiently detect patterns of synchrony between the visual and auditory sensory systems, but also pay more attention to such information than other children. However, as we have noted elsewhere (Falck-Ytter, Rehnberg, & Bölte, 2013) and explain in more detail later on, preference for audiovisual synchrony in the Klin et al. study was assessed using a correlational analysis without control for unisensory (*i.e.*, visual) factors that differed between the stimuli. Therefore, one cannot definitely conclude that the children with autism oriented to multisensory information and not purely visual information. Indeed, when we conducted an experimental study of 3-year-olds with autism where such confounds were avoided, we found the opposite pattern. That is, the children with autism showed no preference for such audiovisual synchrony, while a strong preference was measured for both typically developing 3-year olds and typically developing toddlers (Falck-Ytter et al., 2013). We have also previously tested a similar paradigm in 5-month-old typically developing infants. Even at this young age we found that infants were sensitive to audiovisual synchrony when shown simplified versions of the biological motion stimuli (Falck-Ytter, Bakker, & von Hofsten, 2011).

In sum, there is strong need to clarify the role of multisensory processing early in development in autism in general (Bahrack, 2010; Brandwein et al., 2012; Gogolla et al.,

2014), and to address preferences for audiovisual synchrony expressed within biological motion more specifically (Falck-Ytter et al., 2013; Klin et al., 2009). In the current study, we recruited 10-month-old infant siblings of children with autism (high-risk group) and typically developing controls, and followed them up until three years of age – an age at which autism can be diagnosed reliably. At 10 months of age, processing of biological motion is quite sophisticated in typically developing infants (Booth, Pinto, & Bertenthal, 2002; Kuhlmeier, Troje, & Lee, 2010; Moore, Goodwin, George, Axelsson, & Braddick, 2007; Spencer, O'Brien, Johnston, & Hill, 2006). For example, at this age, infants react with surprise if point-light walkers violate the principle of solidity (Moore et al., 2007). Similarly, infants at this age are sensitive to audiovisual synchrony, in fact, even much younger infants have been shown to be attentive to this basic aspect of their (multimodal) environment (e.g., Bahrick et al., 2004; Spelke, 1979).

In the current study, we experimentally manipulated the spatial distribution of audiovisual synchrony within biological motion while keeping visual information constant, using the same experiment as we have used previously in 3-year-olds with autism (Falck-Ytter et al., 2013). Based on our earlier findings in older children, we here tested the hypothesis that infants with later autism would show reduced orienting to audiovisual synchrony in this context. We expected a reduced preference change both comparing with the typical controls and with high-risk infants who did not fulfil criteria for autism at follow-up.

## **Methods**

### ***Participants***

The participants took part in a longitudinal study (for a general overview, see <http://www.eurosibs.eu/research>; the current eye tracking task was only conducted in Sweden;

www.smasyskon.se). Infants in the HR group were recruited through the project's web site, advertisements, and from clinical units. Infants in the LR group were recruited from the live birth records. Both groups were primarily from the greater Stockholm area. Each infant in the HR group had at least one older sibling with a community diagnosis of ASD. An experienced clinical psychologist interviewed all families upon first contact, and medical records for the older sibling were collected and reviewed to confirm the diagnostic status of the proband. In Sweden, guidelines recommend the use of standardized instruments as part of diagnostic procedures for ASD. In our sample, we could confirm the use of the Autism Diagnostic Observation Schedule (ADOS/ADOS-2) or the Autism Diagnostic Interview – Revised (ADI-R) in 72% of cases through inspection of the obtained medical records. Because details about specific instruments were not always included in the records, the actual figure is likely to be higher.

Written informed consent was collected from parents. The study was approved by the Ethics Board in Stockholm and conducted in accordance with the 1964 Declaration of Helsinki.

Specific exclusion and inclusion criteria followed the consensus in the European longitudinal sibling study EU-AIMS/EUROSIBS ([www.eurosibs.eu/research](http://www.eurosibs.eu/research)). HR group *inclusion*: Older full sibling with autism (presence of community clinical diagnosis, confirmed via inspection of medical records); HR *exclusion*: 1) Diagnosis of epilepsy or history of fits/convulsions in infant, 2) Known presence of genetic syndrome (in proband or infant) clearly related to autism, 3) Presence of known significant uncorrected vision or hearing impairment in infant, 4) Infant was premature (pre 36 weeks), 5) Presence of known significant developmental or medical condition in infant likely to affect brain development or infant's ability to participate in the study. LR group, *inclusion*: Older full-sibling with typical development (by parent

report), exclusion: *Exclusion*: same as above and 1) parent has autism-specific concerns about their infant, 2) Presence of autism in 2<sup>nd</sup> degree relatives

(Table 1 around here)

**Assessments at 10 months of age:** The eye tracking session took place during a full-day visit to the lab, which included several different experiments and assessments. Breaks were included flexibly into program according to the needs of the infant and the parent(s). During eye tracking, infants were seated on their parent's lap, about 60 cm from the eye tracker screen. Eye tracking was conducted using Tobii eye trackers (1750; TX300, Tobii Technology, Danderyd, Sweden). Assessments were preceded by a five-point calibration procedure. Sociodemographic data were collected via online questionnaires, and developmental level was assessed during the same day as the eye tracking, using the Mullen Scale of Early Learning (MSEL, Mullen, 1995) conducted by an experienced clinical child psychologist.

**Assessments at 36 months of age (follow-up):** At 36 months, we collected standardized information on medical history, current developmental and adaptive level as well as problem behaviors using the ADI-R, the ADOS-2 -2, the Vineland Adaptive Behavior Scales, the Child Behavior Checklist 1.5-5, and MSEL during a whole day clinical visit. The clinical evaluation was conducted (without blindness to risk-group status but to the results of the eye tracking task) by experienced clinical researchers (psychologists) with demonstrated research-level reliability. Based on the information, final DSM-5 consensus judgements were made by the clinical researcher together with the last author (SB, international ADI-R, ADOS-2 trainer). Based on this information, participants were assigned either to the autism group, the

high-risk without autism group or the low-risk control group (**Table 1**).

In total, before exclusion for experimental reasons, 34% of the high-risk group received a DSM-5 diagnosis of autism spectrum disorder. This figure is high compared to a large US study (Ozonoff et al., 2011), but similar to the rates previously reported in an European cohort (Elsabbagh et al., 2013). Eye tracking was not part of the 36-month assessment.

### *Experimental stimuli*

Because the design of the stimuli is critical for interpreting findings from studies investigating audiovisual synchrony and might explain discrepancies between studies we will here describe in more detail how we constructed the stimuli used in our previous (Falck-Ytter et al., 2013) and the current report as well as our primary concerns with the previous approaches. The stimuli in the study by Klin et al. displayed point light animations of child-friendly but otherwise diverse actions together with audio (e.g., playing peek-a-boo; enacting a feeding routine). In each stimulus, the same action was always shown in two versions: upright on one side of the screen and spatially inverted (and temporally reversed) on the other (**Figure 1**). Because inversion of point light animations disrupts infants' perception of biological motion (Pavlova & Sokolov, 2000), preferential looking to the upright animation in the pair is commonly used to assess perception of biological motion in infants (Klin et al., 2009; Simion et al., 2008).

(Figure 1 about here)

Although the audio was always in synch with the upright animation, coincidentally synchrony



with the inverted and reverse animation also occurred. Klin et al. (2009) noticed that viewing patterns in children with autism when observing these stimuli could be explained by the distribution of audiovisual synchrony across the two animations in each pair. This was then tested using a correlational analysis, where the distribution of audio-visual synchrony in each video clip was correlated with infants' looking preferences. A concern with this approach is that correlating viewing data across *visually non-identical* stimuli introduces potential confounds. For example, the stimulus may have elicited most orienting to the upright animation not only because of high levels of audiovisual synchrony but also because they displayed a highly repetitive and stereotypic action ("pat-a-cake", i.e., a human actor repeatedly clapping his hands). Fascination for repetitive events and stereotypic behavior are hallmarks of autism; hence it is not unlikely that orienting to the upright biological motion was facilitated in children with autism when viewing actions had these characteristics.

For these reasons, we have developed stimuli that could isolate the effect of audiovisual synchrony and rule out all unisensory factors (Falck-Ytter et al., 2013). Our stimuli resembled the original pat-a-cake stimulus (large arm movements, ~0.65 claps per second, stimulus duration = 15 seconds; see [http://smasyskon.se/biosync\\_stimuli/](http://smasyskon.se/biosync_stimuli/); see also Supplementary Video 2 in (Klin et al., 2009)). As in the original study (Klin et al., 2009), we paired this stimulus with the inverted version of the same animation played in reverse. Moreover, as in the previous study (Klin et al., 2009), both visual stimuli were accompanied by the sound of clapping hands and of a human voice played from a centrally placed loudspeaker. In contrast to the study by Klin et al., we experimentally modified the timing of the auditory clapping to create two visually identical conditions in which the spatial distribution of audiovisual synchrony was reversed. Specifically, in the UPSYNC condition, the auditory clapping was synchronous with visual clapping in the upright animation (and asynchronous with the

clapping in the inverted/reversed animation). This condition is similar to the one used by Klin et al. Conversely, in the INVSYNC condition, the auditory clapping was made synchronous with visual clapping in the inverted/reversed animation. If children with autism indeed orient to audiovisual synchrony as concluded by the authors of the original study (Klin et al., 2009), they should change their preference between the upright and the inverted stimulus as the audio becomes synchronous with one or the other.

When developing the stimuli, we chose a 15-s period from a motion-capture recording (Qualisys, Göteborg, Sweden) which resulted in clearly asynchronous visual clapping between the two animations when one was played backward in time (Falck-Ytter et al., 2011). Since the clapping was one of the main sources of audio stimulation, this maximized the difference in AVS between the upright and inverted stimuli, in both conditions. Examples of the INVSYNC and UPSYNC stimuli are here: [http://smasyskon.se/biosync\\_stimuli/](http://smasyskon.se/biosync_stimuli/)). As can be heard, the soundtrack of the human voice remained unchanged in the two conditions, only the timing of the auditory clapping was changed.

The study also included a third control condition, labelled BIOMOTION, in which the upright and inverted animations used in the UPSYNC/INVSYNC conditions were played forward in time. This caused the magnitude of change in visual motion, and consequently also audiovisual synchrony, to be identical on both sides in each frame of the video. To broadly match the other conditions, the BIOMOTION stimuli included a human voice recording and non-social sounds (no clapping). The purpose of this condition was to be able to test if the groups responded differently to the biological motion present in the other two conditions.

Infants saw four trials of each the UPSYNC and INVSYNC conditions, and eight trials of the

BIOMOTION condition, with left–right counterbalancing of the stimuli. The order of the stimuli (UPSYNC, INVSYNC, BIOMOTION) was randomized.

### *Statistical analyses*

We excluded trials missing more than 50% (7.5 s) of the eye tracking samples, and we excluded infants that did not contribute at least two trials in both the UPSYNC and INVSYNC conditions. This resulted in two infants being excluded in the autism group, eight in the high-risk group without autism and three in the low-risk control group. After this step, there was no significant difference between groups in terms of number of trials included in either condition. We also excluded one participant due to being low-risk control but fulfilling DSM-5 autism spectrum disorder criteria at 3 years. Thus, in total, 14 participants with eye tracking data and outcome data were excluded from the analysis. All participants had at least 2 valid trials in the BIOMOTION condition and the number of trials did not differ between groups.

We were primarily interested in the degree to which infants differentiated between the two experimental conditions, UPSYNC and INVSYNC. Thus, our main dependent measure (**Figure 2**) was a difference score: Preferential looking (%) to the upright animation in the UPSYNC *minus* preferential looking (%) to the upright animation in the INVSYNC condition. The preferential looking data from each condition is reported in **Table 2**. Assuming an effect size identical to the one observed in our earlier study (Falck-Ytter et al., 2013), we had 77% power to detect a difference between the two high risk groups (directional hypothesis).

Visual inspection as well as analyses of the looking data distributions indicated no deviation

from normality (assessed using the Shapiro-Wilk test; all  $P_s > .10$ ) and homogeneity of variance (assessed using the Levene's test; both  $P_s > .1$ ). Moreover, there were no statistical outliers. Therefore, eye tracking data were analyzed with parametric statistical tests with two-tailed probabilities ( $\alpha = .05$ ), unless otherwise specified.

Neither the overall looking time to the screen ( $F(2,44) = .212, P > .25$ ; **Table 2**) nor the amount of looking time in the Area of Interests (AOIs) differed between groups (virtually all data [97%] fell within the AOIs in all groups; hence no inferential statistic is provided for this measure; for definition of AOIs see **Figure 3**). **Figure 3** indicates homogenous performance across groups, in terms of general spatial allocation of gaze.

(table 2 about here)

## Results

As hypothesized, the three groups differed significantly in terms of how much the manipulation of audiovisual synchrony affected their preference for the upright animation in the pair ( $F(2,44) = 4.81, P = .013$ ; **Figure 2**). Furthermore, follow-up analyses (Bonferroni post hoc tests) showed that the autism group differentiated significantly less between the conditions, both when compared to the high risk group without autism ( $P < .032, d = 1.05$ ) and the low risk controls ( $P < .022; d = .96$ ). Both of the two latter groups oriented to audiovisual synchrony ( $P < .001$  and  $P = .01$ , respectively), while the infants with later autism failed to differentiate between the two experimental conditions ( $P > .25$ ). **Table 2** presents looking preference for the upright animation across all groups and conditions; **Figure 3** shows the same data represented as heatmaps superimposed on the stimuli.

(Figure 2 about here)

Because audiovisual synchrony was expressed within biological motion in the two experimental conditions, we wanted to check whether the groups may have reacted differently to the biological motion *per se*. However, we found no indication that the groups differed in terms of preference for the upright animation in the BIOMOTION stimuli ( $F(2,44)=0.142$ ,  $P>.25$ ; autism: 52% [sd=10%], control groups: 52% [sd=8%] and 53% [sd=8%], respectively; see also **Table 2**). These negative findings should, however, be interpreted with caution given that we had only 64-70% power to detect large effect sizes (0.8) for specific group comparisons (one-tailed, based on the hypothesis that biological motion preference, if anything, should be reduced in autism; Klin et al., 2009). Nevertheless, the data from the BIOMOTION control condition suggest that that the main results (**Figure 2**) are not attributable to differential responding to the visual information in the stimuli *per se*. As expected, overall the infants in the study showed a significant preference for the upright animation in the BIOMOTION condition ( $t(46)=2.01$ ,  $P=.047$ , one sample t-test,  $d=.29$ ).

Autism primarily affects males, but our autism group included a relatively high proportion of female infants (**Table 1**). Although we have no a priori reason to expect the girls in the sample drove our results, it would be a problem for the generalizability of our results to the general autism population, if that was the case. In order to investigate this issue, we checked if the main results (**Figure 2**) still hold if girls were excluded from analysis, keeping in mind that this results in a much smaller sample size. Indeed, the magnitude of difference in looking time between the two conditions (UPSYNC, INVSYNC) remained significantly different between the three groups ( $F(2,20)=5.707$ ,  $P=.011$ ). Planned comparisons showed that boys with later autism differentiated less between the conditions both compared with boys in the

high risk group without autism ( $P=.025$ ,  $d=1.49$ ) and with boys in the typical control group ( $P=.004$ ,  $d=1.54$ ). The direction of differences were the same for girls, but not significant.

This shows that the main result (**Figure 2**) is not attributable to the relatively high proportion of girls in our sample.

(Figure 3 about here)

## **Discussion**

As predicted, we found that the infant siblings who later received an autism diagnosis were quantifiably less inclined to orient to audiovisual synchrony when expressed in point light displays of biological motion compared to two groups of infants who did not fulfill criteria for autism at follow up. Thus, the finding from the current longitudinal study of infants at risk does not support the view that excessive orienting to audiovisual synchrony in social contexts plays an important role in shaping the developmental trajectory in infants at risk for autism, as was suggested earlier. Specifically, based on their findings, Klin et al. (2009) concluded that “By two-years-of-age, [...] children [with autism] are on a substantially different developmental course, having learned already from a world in which the physical contingencies of coincident light and sound are quantifiably more salient than the rich social information imparted by biological motion.” (p 260). Rather, if anything, the current study suggests that *reduced* orienting to synchrony could be play a role in the early development of this group. More generally, our result is consistent with the idea that infants with later autism may process multisensory information less efficiently than infants who do not go on to develop autism (Bahrick, 2010; Brandwein et al., 2012; Falck-Ytter et al., 2013; Gogolla et al., 2014). From a theoretical point of view, it will be important to study the relation between

multisensory processing and brain connectivity (Emerson et al., 2017; Wolff et al., 2012), as well as how it relates to excitation-inhibition imbalance in developing brain networks in humans (Gogolla et al., 2014; Hattori et al., 2017). More insight into multisensory processing difficulties in autism could also have practical implications in the longer run, particularly in light of recent evidence that very brief training leads to lasting improved audiovisual temporal processing in adults (Powers, Hillock, & Wallace, 2009).

Given that the groups performed similarly in terms of their preference for the upright animation in the BIOMOTION control condition, the reduced orienting to audiovisual synchrony seen in the autism group cannot be explained by atypical orienting to the biological motion information in these particular stimuli. Descriptively, across all three conditions, a similar viewing preference for the upright animation was observed in the autism group (**Table 2**). Although the result tentatively speaks against the hypothesis that reduced orienting to biological motion is an important factor in the emergence of autism (Falck-Ytter et al., 2013; Klin et al., 2009), more studies are needed to evaluate this issue. Although the preference for the biological motion stimulus was significant in the total sample, the magnitude of this effect was small. Previous research has shown age-related changes in preference for biological motion in infancy (Fox & McDaniel, 1982). Moreover, the magnitude of infants' preference for biological motion in the paired-visual preference paradigm is likely to depend on the particular stimuli characteristics, and generalization from one stimulus to another may not be straightforward. This motivates systematic longitudinal studies assessing biological motion processing in different contexts, and using brain as well as behavioral measures.

Given that the results of the current and our previous reports differ substantially from the results of the Klin et al. (2009) study, a more detailed analysis of this discrepancy is

warranted. First, as we have argued above (see Introduction and Methods), due to the correlational design of their study, Klin et al. were not able to rule out potential unisensory confounds, such as the presence of highly repetitive movement in the stimulus that elicits the strongest preference in the autism group. While our studies, both the current and the previous one (Falck-Ytter et al., 2013), fail to support the general conclusion by Klin et al., it is important to note that our studies are not direct replication attempts, but rather attempts to address the conceptual questions with improved methodology. This leads to the next important point: Do our experimental conditions represent a valid conceptual replication of the earlier work? To address this, it is necessary to consider the UPSYNC and INVSYNC conditions in more detail. The UPSYNC condition in the current study was very similar to the stimulus in the original report by Klin et al. (2009) that caused the strongest preference for audiovisual synchrony in their autism group. Indeed, we replicate the earlier results for this particular condition in the sense that the performance of the autism group and the low-risk control group was not significantly different. The INVSYNC condition (for an illustration, see [http://smasyskon.se/biosync\\_stimuli/](http://smasyskon.se/biosync_stimuli/)) on the other hand, was the result of our experimental manipulation and had no analogy in the earlier study (audiovisual synchrony was not selectively manipulated in their study; the soundtrack was always the one naturally accompanying the upright animation). One potential criticism could therefore be that our INVSYNC condition was unnatural, because we manipulated the audio track so that the sound of the clapping was out of synchrony with the visual clapping in the upright animation and in synchrony with the inverted and reversed animation. However, as shown in Figure 3 in the original study (Klin et al., 2009), the proposed model for how audiovisual synchrony affected viewing in autism did not distinguish between audiovisual synchrony expressed in the upright (“natural”) or the inverted/reversed (“unnatural”) animation. The audiovisual synchrony in the inverted animation in the original study occurred entirely by chance alignment of the



audio signal (played forward) and the visual animation (played backward), hence the audiovisual synchrony from this animation was clearly not natural in any sense. For some of the clips used in that study (see e.g. Fig 3h and Figure 3 in (Klin et al., 2009), a large percentage of the audiovisual synchrony was expressed in the inverted/reversed animation, and this synchrony was taken into account to predict looking preference. Thus, an argument stating that the current INVSYNC condition included unnatural audiovisual synchrony would equally apply to our study and to Klin et al. Second, it is possible that the discrepancy between studies is due to age differences (or other differences in sample characteristics), as we have studied 3-year-olds and 10-month-olds with (later) autism, while Klin et al. studied 2-year-olds. Previous “infant sibs” studies have described developmental changes in the manifestation of various autism antecedents, mainly occurring within the first two years of age (e.g. Jones & Klin, 2013; Wolff et al., 2012), but not yet a double reversal as would be the case if we need to accommodate both our findings and those from Klin et al.

Several limitations of the current study should be mentioned. First, the study was designed to test a specific hypothesis about infants’ preference for audiovisual synchrony within biological motion– and follow up studies are needed to clarify underlying processes as well as the generalizability of the group differences to other contexts. Second, as the current groups were relatively small, independent replication is important. Sample size is a general issue in sibling studies, and our sample size is in line with several previous reports (e.g., Elsabbagh et al., 2012; Jones & Klin, 2013). The main result (**Figure 2**) in our study was based on a directional hypothesis which we had reasonable power to detect. As noted above, whether the negative findings will generalize is more questionable, and this needs to be addressed in future work. Third, clinicians were not unaware of risk group, which may have biased diagnostic outcome determination. Importantly, however, the clinicians were unaware of eye tracking results; hence this bias cannot explain the observed difference within the high

risk group (**Figure 2**). Finally, there was a higher than expected rate of autism outcomes (34%) in this high-risk group, as well as an unexpectedly high proportion of females in the autism group (38% versus the typical rate of 20%). The fact that we find that the results replicate when boys only are considered, indicates that the main result nevertheless would generalize to the typical autism population (consisting mainly of boys).

## **Conclusion**

In sum, our findings suggest that infants with typical development, but not infants with later autism, orient to audiovisual synchrony within biological motion. This conclusion is opposite to the one drawn in the study of two-year-olds (Klin et al., 2009), but in line with our earlier studies of three-year-olds with autism using the current experimental paradigm (Falck-Ytter et al., 2013). The finding that typically developing infants orient to audiovisual synchrony is in line with a large literature on the adaptive value of orienting to multisensory information in typical development (Bahrick et al., 2004). The current results support domain general (“non-social”) accounts of autism (Elsabbagh & Johnson, 2016; Hazlett et al., 2017) and the poor multisensory processing hypothesis more specifically (Balz et al., 2016; Brandwein et al., 2012; Gogolla et al., 2014; Stevenson et al., 2014). The findings, particularly if they generalize to multisensory processing more broadly, could have far-reaching consequences for our understanding of early developmental pathways in autism, and potential implications for early intervention (Powers et al., 2009). Our findings also highlight the need for independent replication of results in the “early autism” field.

## **Key Points**

- Multisensory processing difficulties has been found in individuals diagnosed with autism
- No study has investigated whether such atypicalities could play a role in the early development of the disorder
- We find reduced orienting to audiovisual synchrony in infants who later receive an autism diagnosis
- This finding supports the view that autism is linked to atypical multisensory processing early in life
- This result has potential implications for early detection and intervention

## **Acknowledgements**

The EASE team includes: Sheila Achermann; Linn Andersson Konke; Karin Brocki; Elodie Cauvet; Martina Hedenius, Johan Lundin Kleberg; Elisabeth Nilsson Jobs, Emilia Thorup.

This study was supported by seed funding from Uppsala University (TFY), Stiftelsen Riksbankens Jubileumsfond (P12-0270:1; NHS14-1802:1; Pro Futura Scientia [in collaboration with the Swedish Collegium for Advanced Study, SCAS]), the Swedish Research Council (2015–03670; 523-2009-7054), EU (MSC ITN 642996), the Strategic Research Area Neuroscience at Karolinska Institutet (StratNeuro), as well as from the Swedish Research Council for Health, Working Life and Welfare in collaboration with Swedish Research Council, FORMAS and VINNOVA (259-2012-24). TG is funded by the Medical Research Council Program Grant G0701484. The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115300, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007 - 2013) and EFPIA companies' in kind contribution.

**Correspondence to:**

Dr Terje Falck-Ytter

Department of Psychology  
Box 1225, 751 42 Uppsala  
Uppsala University  
Sweden

Phone: +46 (0) 70 4581475

Email: [terje.falck-ytter@psyk.uu.se](mailto:terje.falck-ytter@psyk.uu.se)

## References

- Bahrick, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. In J. G. Bremner & T. D. Wachs (Eds.), *Blackwell handbook of infant development*: Wiley-Blackwell.
- Bahrick, L. E., Liekliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions In Psychological Science*, 13(3), 99-102.
- Balz, J., Keil, J., Romero, Y. R., Mekle, R., Schubert, F., Aydin, S., . . . Senkowski, D. (2016). GABA concentration in superior temporal sulcus predicts gamma power and perception in the sound-induced flash illusion. *Neuroimage*, 125, 724-730.
- Booth, A. E., Pinto, J., & Bertenthal, B. I. (2002). Perception of the symmetrical patterning of human gait by infants. *Developmental Psychology*, 38(4), 554.
- Brandwein, A. B., Foxe, J. J., Butler, J. S., Russo, N. N., Alschuler, T. S., Gomes, H., & Molholm, S. (2012). The development of multisensory integration in high-functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cerebral Cortex*, bhs109.
- Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T., Johnson, M. H., & British Autism Study Infant, S. (2013). Disengagement of Visual Attention in Infancy is Associated with Emerging Autism in Toddlerhood. *Biological Psychiatry*, 74(3), 189-194. doi:10.1016/j.biopsych.2012.11.030
- Elsabbagh, M., & Johnson, M. H. (2016). Autism and the social brain: the first year puzzle. *Biological Psychiatry*.
- Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., . . . Team, B. (2012). Infant Neural Sensitivity to Dynamic Eye Gaze Is Associated with Later

- Emerging Autism. *Current Biology*, 22(4), 338-342. doi:10.1016/j.cub.2011.12.056
- Emerson, R. W., Adams, C., Nishino, T., Hazlett, H. C., Wolff, J. J., Zwaigenbaum, L., . . . Piven, J. (2017). Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Science Translational Medicine*, 9(393). doi:10.1126/scitranslmed.aag2882
- Falck-Ytter, T., Bakker, M., & von Hofsten, C. (2011). Human infants orient to biological motion rather than audiovisual synchrony. *Neuropsychologia*, 49(7), 2131-2135.
- Falck-Ytter, T., Rehnberg, E., & Bölte, S. (2013). Lack of Visual Orienting to Biological Motion and Audiovisual Synchrony in 3-Year-Olds with Autism. *Plos One*, 8(7). doi:e6881610.1371/journal.pone.0068816
- Fox, R., & McDaniel, C. (1982). The Perception Of Biological Motion By Human Infants. *Science*, 218(4571), 486-487.
- Gogolla, N., Takesian, Anne E., Feng, G., Fagiolini, M., & Hensch, Takao K. (2014). Sensory Integration in Mouse Insular Cortex Reflects GABA Circuit Maturation. *Neuron*, 83(4), 894-905. doi:10.1016/j.neuron.2014.06.033
- Hattori, R., Südhof, T. C., Yamakawa, K., & Hensch, T. K. (2017). Enhanced cross-modal activation of sensory cortex in mouse models of autism. *bioRxiv*. doi:10.1101/150839
- Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., . . . Botteron, K. N. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, 542(7641), 348-351.
- Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, 504(7480), 427-+. doi:10.1038/nature12715
- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*,

459(7244), 257-261.

Kuhlmeier, V. A., Troje, N. F., & Lee, V. (2010). Young infants detect the direction of biological motion in point-light displays. *Infancy*, *15*(1), 83-93.

Moore, D. G., Goodwin, J. E., George, R., Axelsson, E. L., & Braddick, F. M. (2007). Infants perceive human point-light displays as solid forms. *Cognition*, *104*(2), 377-396.

Mullen, E. M. (1995). *Mullen Scales of Early Learning: AGS Edition*. MN: AGS: Circle Pines.

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., . . . Stone, W. L. (2011). Recurrence Risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium Study. *Pediatrics*, *128*(3), E488-E495.  
doi:10.1542/peds.2010-2825

Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*(5), 889-899.

Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of Neuroscience*, *29*(39), 12265-12274.

Russo, N., Foxe, J. J., Brandwein, A. B., Altschuler, T., Gomes, H., & Molholm, S. (2010). Multisensory processing in children with autism: high-density electrical mapping of auditory–somatosensory integration. *Autism Research*, *3*(5), 253-267.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *105*(2), 809-813.

Spelke, E. S. (1979). Perceiving Bimodally Specified Events In Infancy. *Developmental Psychology*, *15*(6), 626.

Spencer, J., O'Brien, J., Johnston, A., & Hill, H. (2006). Infants' discrimination of faces by



using biological motion cues. *Perception*, 35(1), 79.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: The MIT press.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G.,

Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, 34(3), 691-697.

Wolff, J. J., Gu, H., Gerig, G., Elison, J. T., Styner, M., & Gouttard, S. (2012). Differences in White Matter Fiber Tract Development Present From 6 to 24 Months in Infants With Autism. *The American Journal of Psychiatry*, 169(6), 589-600.

doi:10.1176/appi.ajp.2011.11091447

**Tables****Table 1**  
**Participant characteristics (final sample; 10 months)**

	Low risk controls (7/14 females)				High risk without autism (12/20 females)				High risk with autism (5/13 females)				Group comparison
	N	Min	Max	Mean	N	Min	Max	Mean	N	Min	Max	Mean	Kruskal Wallis Test
Age (days)	14	296	343	313.8	20	276	332	308.8	13	292	364	313.4	p>.25
Education	14	3	5	4.7	20	2	5	4.3	13	1	5	3.9	p=.13
Family income	13	3	7	5.1	18	2	7	4.7	13	2	6	4.2	p>.25
MSEL_VR	14	8	14	11.4	20	7	17	11.3	13	8	13	10.8	p>.25
MSEL_FM	14	8	16	12.1	20	7	15	11.7	13	10	16	12.7	p>.25
MSEL_RL	14	6	14	9.1	20	5	14	9.9	13	6	13	9.2	p>.25
MSEL_EL	14	5	13	9.1	20	6	13	9.1	13	4	13	8.9	p>.25
MSEL_total	14	82	126	102.0	20	79	128	102.6	13	78	121	100.6	p>.25

**Notes:**

Education is measured on a 5-point scale, where 5 represents higher university education >3 years.

Family income is based on a 7-point scale, where 7 represents a family income of 70,000 SEK/month or higher

MSEL\_VR=Mullen scales of early learning - visual reception subscale

MSEL\_FM= fine motor subscale

MSEL\_RL= receptive language subscale

MSEL\_EL= expressive language subscale

MSEL\_total=Mullen scales of early learning - total score

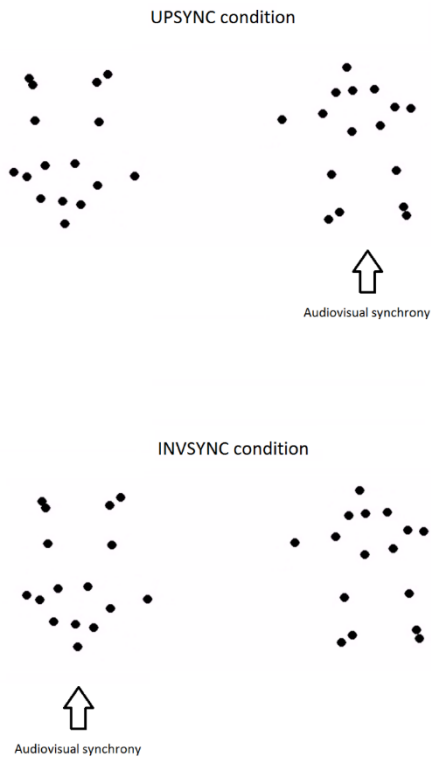
Most measures did not fulfill criteria for parametric testing, hence we used Kruskal Wallis throughout for simplicity

**Table 2**  
**Eye tracking data – descriptive statistics**

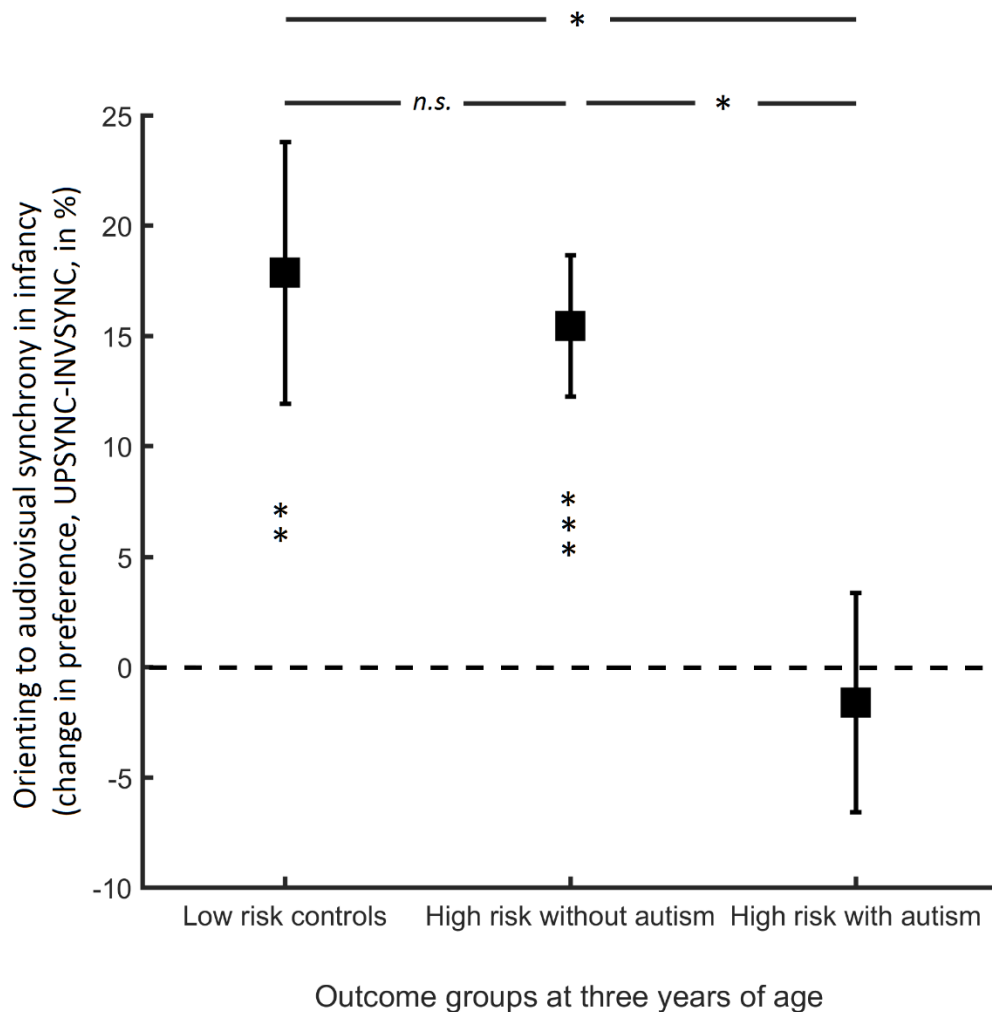
		Low risk controls		High risk without autism		High risk with autism	
		MEAN	SD	MEAN	SD	MEAN	SD
Overall looking time (%)	Screen relative to full stimulus duration	84	7	85	10	83	9
	AOIs relative to screen	97	1	98	1	97	1
Valid Trials (%)	UPSYNC	77	21	88	19	77	22
	INVSYNC	77	22	86	17	79	22
	BIOMOTION	73	24	88	19	72	24
Preference for upright animation (%) <sup>1</sup>	UPSYNC	59	14	67	17	55	18
	INVSYNC	41	17	52	11	57	11
	BIOMOTION	52	9	53	9	52	10

<sup>1</sup>(Upright AOI/(Upright AOI+Inverted AOI))\*100

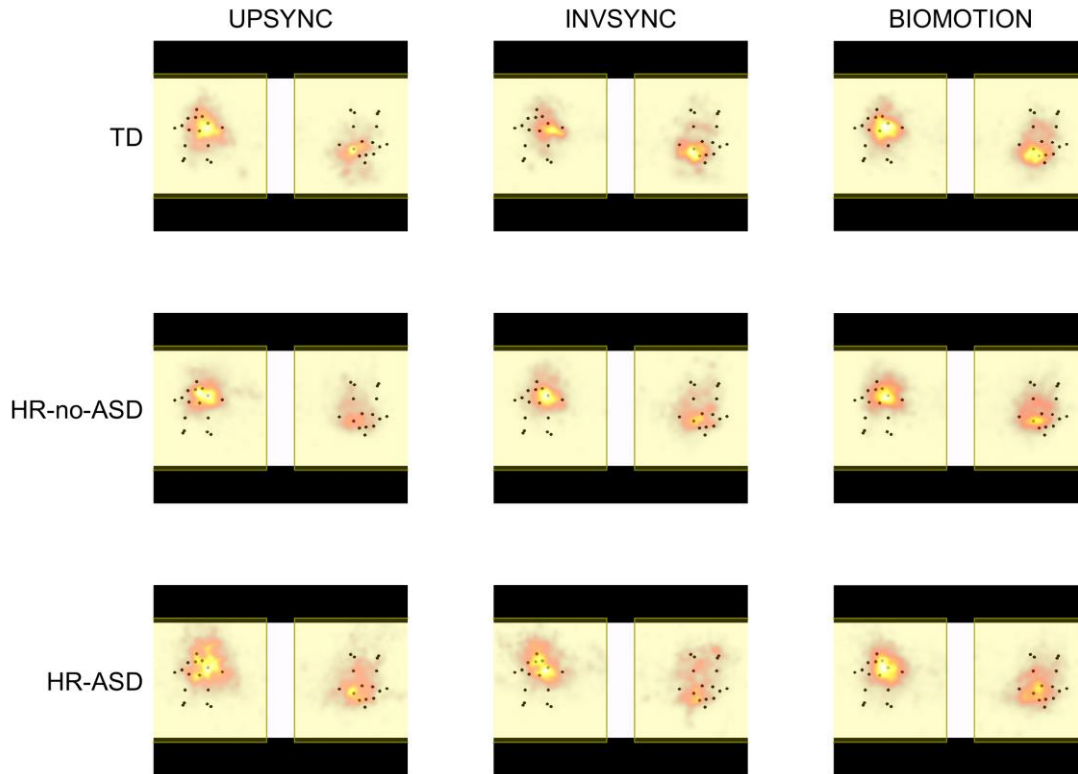
## Figure legends



**Figure 1.** Experimental stimuli. An upright animation (right) of a person clapping hands was shown side-by-side with a spatially inverted (upside-down) and temporally reversed version of the same animation (left). The visual stimuli were accompanied by an audio track including a child friendly human voice and the sound of clapping hands. In the UPSYNC condition (top) the auditory clapping was synchronous with the visual clapping in the upright animation. In the INVSYNC condition (bottom), the auditory clapping was synchronous with the visual clapping in the inverted animation. See main text for further information.



**Figure 2.** Reduced orienting to audiovisual synchrony in infancy predicts autism at three years of age. We calculated preferential looking (%) to the upright animation in the UPSYNC and INVSYNC conditions, and subtracted the latter percentage from the former. This difference score is shown on the y-axis; positive values indicate that one looks more to the upright animation when that animation is in synchrony with the audio than when the other (inverted) animation is in synchrony with the audio. Only the two non-autistic groups changed their viewing preference because of the experimental manipulation. See also **Table 2**. Error bars = s.e.m.; \* $P < .05$  (Bonferroni post hoc test); \*\* $P = .01$ ; \*\*\* $P < .001$ .



**Figure 3.** Gaze heatmaps for each group and condition. Infants in the three groups attended to roughly the same areas, suggesting that the main results are not attributable to atypicalities in attention to the displays in autism. Across all groups and conditions, the average number of gaze shifts per trial fell in the range 3.01 (sd=1.01) to 4.40 (sd=2.34), again suggesting similar scanning of the stimuli. TD: Typically developing; HR-no-autism: high-risk children without autism; HR-autism: High-risk with autism. Areas of Interest (AOIs) used for analysis are superimposed (yellow squares).