33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

# Measuring Gaze of Children with Autism Spectrum Disorders in Naturalistic Interactions

Basilio Noris, Mandy Barker, Jacqueline Nadel, François Hentsch, François Ansermet and Aude Billard

*Abstract*— **In this paper, we report on a study on gaze behavior by children with Autism Spectrum Disorder (ASD) during a dyadic interaction in a naturalistic environment. Twelve children with ASD were contrasted to twelve typically developing (TD) children, in a semi-structured interaction with a selection of items from the Early Social Communication Scale (ESCS). We used the WearCam, a novel head-mounted eye-tracker designed for children, to obtain gaze information across the broad field of view from the viewpoint of the child. Children with ASD looked downwards more often, and explored their lateral field of view more extensively compared to TD children. We discuss a number of hypotheses in support of these observations.**

## I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a developmental disorder that touches 1 every 160 children [1]. ASD presents itself in impairments in social interaction, communication and in stereotypical and repetitive behaviors. A number of atypical visual behaviors and viewing strategies have been reported in children with autism, such as reduced gaze towards social stimuli, preference for the mouth region rather than the eyes when looking at a face or difficulties in attention shifts [2], [3], [4]. The study of how children explore and scan their field of view, and how this relates to their way of observing or interacting visually with objects and people, has been the subject of some work. Early studies of children with ASD observed that children often *looked from the corner of the eyes* [5] when looking at other people. A symptom commonly associated with ASD is the phenomenon of downcast gaze [6]. In a study of children with ASD playing and exploring objects, [7] reported an unusual amount of *lateral glances*, or instances of gaze directed to the extreme of the lateral (horizontal) field of view. These observation have been linked to accounts of hypo- and hyper-sensitivity to visual stimuli, and suggest the use of these atypical visual behavior as compensating strategies [7]. The quantitative evidence in support of these hypotheses is, however, still scarce [8].

The studies on visual behavior in ASD are usually conducted either in laboratory settings, using monitor-based eye-trackers ([3], [4] to give a few examples), or in naturalistic settings, in which observation through external video cameras is manually analyzed. The former approach allows to automatically obtain precise measurements, but the interaction is limited as the tracking is restricted in the space of a monitor. The latter requires an extremely long analysis time, as multiple raters have to manually assess the behavior of the child to ensure reliable results (e.g. [7]). However, naturalistic interactions can elicit more natural behaviors from the child [9], [8], and can therefore provide a better insight on the usual behavior of individuals with autism.

To address the drawbacks of evaluation in naturalistic interactions, we have developed the WearCam [10], a head-mounted eye-tracker designed specifically for young children, which allows the recording of the field of view as seen by the child, and to simultaneously record the direction of the child's gaze. The technical aspects of the device (shown in Figure 1) will be described in Section II-B. The WearCam is novel in that it allows to study the gaze of the child during natural interactions from the point of view of the child, and this across the whole field of view. This opens the door to the measurement of elements of visual behavior, such as gaze directed to the side of the field of view, that have not been studied extensively yet.

In this paper we present a study of 24 children (12 children with ASD, and 12 typically developing) interacting in a natural environment with an adult experimenter. We have chosen a selection of items from the Early Social Communication Scale (ESCS [11]), to display the potential of using the WearCam in the scope of clinical assessment and research. We focus our analysis on the use of gaze across the whole field of view.

Fig. 1. *Left*: The WearCam worn by a typically developing 14 months old girl, fixed on a cap. *Right*: Closeup of the WearCam, with the two cameras and motor-controlled mirror.

B. Noris and A. Billard H. are with the Learning Algorithms and Systems Laboratory, Ecole Polytechnique Fdrale de Lausanne (EPFL), Switzerland basilio.noris@epfl.ch

M. Barker is with the Department of Child and Adolescent Psychiatry, Centre Hospitalier Universitaire de Lausanne (CHUV), Switzerland

J. Nadel is with the CNRS Emotion Center, Université Pierre et Marie Curie, Pitié-Salpétrière, France

F. Hentsch and F. Ansermet are with the Department of Child and Adolescent Psychiatry (SPEA), School of Medicine, University of Geneva, Switzerland

978-1-4244-4122-8/11/$26.00 ©2011 IEEE 

## II. METHODS

### A. Participants

We recruited twelve children with ASD (11 boys, 1 girl) from the Child Psychiatric departments of the University hospitals of Lausanne and Geneva in Switzerland. Their Chronological Age (CA) in years was $[2.9-8.8]$ (M=5.3 SD=1.8). Their diagnosis was confirmed using the ADI-R test. They were matched on gender and Developmental Age (DA) with twelve Typically Developing (TD) children (CA: $[1.1-7.1]$ M=3.3 SD=1.9). The DA were assessed using the Vineland Adaptive Behavior Scale [12]. The mean scores and SD of the two groups for each sub-scale are presented on Table I. Each child participated to a recording session that lasted a maximum of 10 minutes. All parents had given their written informed consent including permission to use video recordings and pictures of the children for scientific publications. The experimental protocol and consent form had been approved by the Ethics Committee of the University Hospitals of Geneva and Canton de Vaud.

TABLE I
SCORES OF THE ASD AND TD CHILDREN ON THE VINELAND ADAPTIVE BEHAVIOR SCALES

| Variable | ASD$^a$ | TD$^a$ |
|---|---|---|
| Communication | $2.7 \pm 2.3$ | $3.0 \pm 2.0$ |
| Autonomy | $2.8 \pm 1.9$ | $2.7 \pm 1.8$ |
| Socialization | $2.0 \pm 1.5$ | $2.7 \pm 1.7$ |
| Mobility | $3.2 \pm 1.3$ | $2.3 \pm 1.3$ |
| **Adaptive Behavior**$^b$ | $\mathbf{2.9 \pm 1.7}$ | $\mathbf{2.9 \pm 1.6}$ |
| Chronological Age | $5.3 \pm 1.8$ | $3.3 \pm 1.9$ |

$^a$ *mean ± standard deviation*
$^b$ *used to match devel. age*

### B. Apparatus

The experiments were recorded using the WearCam [10], a head-mounted eye-tracker designed specifically for children. The system is composed of two cameras placed on the forehead of the child, which record the field of view in front of the child, and of a small mirror, driven by a servo motor allowing to reflect the eyes of the child into one of the two cameras (see Figure 2). The WearCam is mounted on an adjustable strap or on a child's cap, and weights 180g (straps/cap included). its cameras record a combined image of 384x576 pixels at 25 or 50 frames per second. The field of view recorded by the WearCam is very wide (96° by 96°), which is more than twice that of any existing eye-tracker, and allows to study gaze when it is not directed in the middle of the field of view. A mapping from the pixel information of the eyes image to the coordinates of the gaze direction is obtained through Support Vector Regression [13], which allows to extract information about gaze without explicitly computing geometrical features of the eyes. This allows the system to compute the direction of gaze even when the iris and pupil are not visible (e.g. when the child is looking downwards) and therefore to have a coverage of gaze tracking across the whole field of view from the WearCam images. The SVR mapping from the eyes image allows to compute the direction of the gaze of the child with an average accuracy of 2.42° [10]. The WearCam does not require an active calibration from the child: a calibration phase is performed by the experimenter post hoc; this facilitates the recording phase and allows the equipment to be used with children who might not be able to cooperate in a calibration phase.

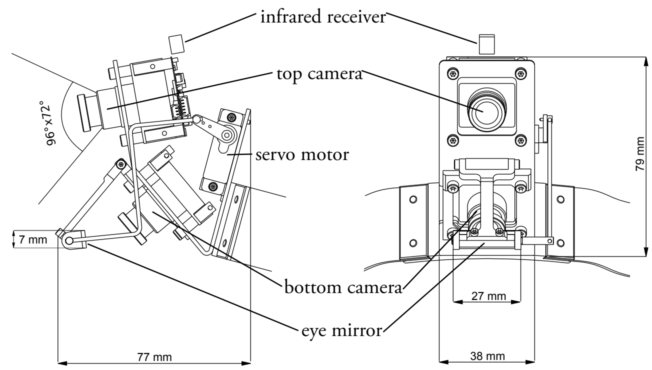


Fig. 2. Schematics of the WearCam, with the two CCD cameras, servo motor and mirror reflecting the eyes of the child.

### C. Procedure

The experimental protocol comprised four items selected from the abridged version of the ESCS[1] test [11]. The items presented were: *Blowing soap bubbles*, *Playing with a wind-up mechanical toy*, *Playing with a toy car* and *Playing with a small ball*. These items focused on the observation and exploration of objects (Soap Bubbles and Wind-up Toy) and on turn taking (Small Ball and Toy Car). The protocol lasted in all cases between 5 and 10 minutes and was administered in a naturally lit room. The child was sitting at a table on a child-sized chair, while the experimenter administering the protocol sat at the opposite side of the table also on a low chair (see Figure 3). The experimenter presented the items and interacted with the child. At all times, the people present in the room consisted of the child, the experimenter, a silent observer and a parent. The parent was placed behind the child and did not interact with her for the duration of the experiment. The observer also was placed behind the child at a distance of several meters so as to minimize the interference on the child's attention.

As the WearCam requires no calibration, the experiment started as soon as the WearCam had been fastened to the child's head and the mirror aligned so that the child's eyes were clearly visible in the camera's image. In a few instances, the camera moved on the head of the child during the experiment (5 instances out of 20 recordings). When that happened, the observer would use a remote control to realign the mirror with the eyes of the child. These occurrences did

[1]The ESCS clinical test is a 20-minute videotaped structured observation that enables assessment of a child's initiation and response to nonverbal communication acts (joint attention, social interaction behaviors, requesting behaviors). The ECSC is administered routinely at the CHUV/HUG during clinical screening of ASD in nonverbal children.

not seem to interrupt the experiments and did not distract the child.

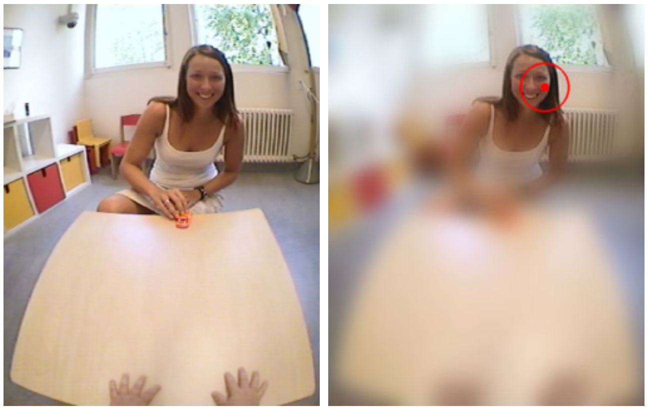

Fig. 3. Image of the interaction with the experimenter recorded from the WearCam. Gaze direction extracted by the eyes image; the region outside central vision (10° radius around the gaze point) has been blurred to highlight the direction of gaze.

### D. Data Analysis

The analysis of the data from the WearCam was obtained through the algorithms described in the previous section. Three trained raters collected calibration samples for the gaze tracking for each recording individually. The collection of calibration samples was effectuated as follows. The rater opened the video in a custom-made software, which displayed the original image of the WearCam, an image of the child's eyes and an image of the field of view recomposed from the two camera images. The rater went through the video and identified instances in which it was easy to infer the direction of the gaze from the image of the eyes and the context (e.g. when the child reached for an object, the eyes shifted in the direction of the object, which allowed the experimenter to accurately place a calibration point on the object). The system automatically recalibrated for each new calibration sample. The calibration procedure stopped when 50-70 calibration samples had been collected. This process took on 10-15 minutes per video. The raters were blind to the goals of the study and to the diagnosis of the participants. We assessed the inter-rater reliability on the tracking of gaze, by comparing the rating of over 40 minutes of video that was labelled by all three raters. The correlation between raters was $> 0.9$.

To maintain consistency across experiments, each recording was split into four parts corresponding to each item presented to the child. These corresponded to each ESCS item presented by the psychologist. This segmentation resulted in item-subsets of durations ranging from 1 to 3 minutes. For each recording, we computed the trajectories of gaze across the broad field of view, and evaluated their distribution (mean angles and standard deviations) across the vertical and lateral axes. For each child, an average of 15.95 min (SD=5.26) of usable video were collected (this corresponds to $\sim$ 24'000 (SD 7'800) frames in which gaze information was present).

We collected the trajectories of gaze for all the recordings and combined the coordinates of gaze from each group to obtain two histograms of the gaze direction throughout the experiments. We then defined the following measurement variables *Gaze Vertical Mean* ($X_1$), *Gaze Vertical Dispersion* ($X_2$), *Gaze Lateral Mean* ($X_3$), and *Gaze Lateral Dispersion* ($X_4$). Dispersion was computed as the standard deviation of the gaze distribution. A mixed design ANOVA test was run independently for $X_{1,\dots,4}$ with between-subject factor *Diagnosis* ({ASD, TD}) and within-subject factors *Developmental Age* (Continuous variable) and *Item* ({bubbles, mouse, car, ball}). We verified the gaussianity of the distribution of our measurements using a Kolmogorov-Smirnoff test, and ran student t-tests on each measured variable accounting for the Diagnosis factor.

TABLE II

3-WAY ANOVAS ON THE FACTORS VERTICAL MEAN AND LATERAL DISPERSION

| Source | Sum Sq. | d.f. | Mean Sq. | F | P |
|---|---|---|---|---|---|
| | | Vertical Mean | | | |
| **Autism** | 1177.12 | 1 | 1177.12 | **12.82** | **0.0008** |
| DevAge | 0.34 | 1 | 0.34 | 0 | 0.9518 |
| **Item** | 1989.19 | 3 | 663.06 | **7.22** | **0.0004** |
| * Total | 7993.96 | 87 | | | |
| | | Lateral Dispersion | | | |
| **Autism** | 72.71 | 1 | 72.707 | **5.94** | **0.0185** |
| DevAge | 42.34 | 1 | 42.3448 | 3.46 | 0.0689 |
| **Item** | 156.71 | 3 | 52.2361 | **4.27** | **0.0094** |
| * Total | 1037.54 | 87 | | | |

*Interactions and error omitted for space purposes.

## III. RESULTS

The results collected from the analysis of the recordings is summarized in Table II and III. The 3-Way ANOVA on all variables revealed a main effect on the $X_1$ and $X_2$ variables for *Diagnosis* ($F_{X_1}(1,87) = 12.82, P < 0.001$ and $F_{X_2}(1,87) = 5.94, P = 0.019$) and *Item* ($F_{X_1}(3,87) = 7.22, P < 0.001$ and $F_{X_2}(3,87) = 4.27, P = 0.009$). No interaction between factors was observed. The effect of item on the child's gaze is not surprising, as different tasks may elicit different types of gaze behavior (e.g. blowing bubbles vs. playing with a toy car). As each child played each item for comparable amounts of time, the experiment is not biased by this effect. Figure 4 displays a boxplot representation of the use of Vertical and Lateral gaze. Compared to TD children, children with ASD presented gaze patterns that appeared more often in the lower part of the vertical field of view, and that were more dispersed in the lateral field of view. Indeed, children with ASD kept their gaze directed significantly lower than TD children ($-0.66°$ vs $7.63°$, $p < 0.001$, DF:86) and explored the lateral field of view more extensively ($13.74°$ vs $11.47°$, $p < 0.005$, DF:86).

TABLE III

STUDENT T-TESTS ON GAZE VARIABLES

| | TD | ASD | t-test |
|---|---|---|---|
| **Vert Mean** | $7.6° \pm 11.4°$ | $-0.7° \pm 9.5°$ | **p: 0.000 (DF: 86)** |
| Vert Disp. | $15.3° \pm 3.9°$ | $15.5° \pm 4.7°$ | p: 0.816 (DF: 86) |
| Lat Mean | $1.1° \pm 7.4°$ | $-0.2° \pm 7.6°$ | p: 0.323 (DF: 86) |
| **Lat Disp.** | $11.5° \pm 3.4°$ | $13.7° \pm 4.9°$ | **p: 0.005 (DF: 86)** |

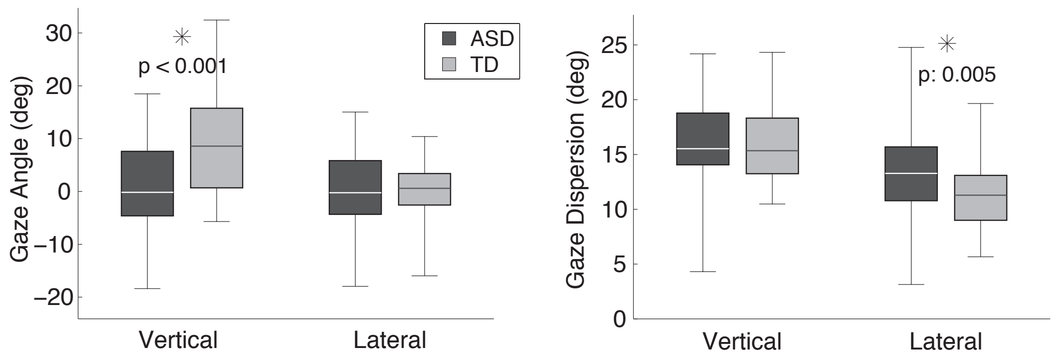


Fig. 4. Mean vertical and lateral angles of the gaze *(left)* and exploration of the gaze in the vertical and lateral directions *(right)*

## IV. DISCUSSION

Our results are in support of the phenomenon of downcast gaze in autism. [6] explains the downcast gaze as a response sensory overload coming from a hypersensitivity to visual stimuli. By looking downward, children with ASD can look at static elements (table, ground) that are less likely to perturb them. In our experiments, the upper part of the field of view (see Figure 3) contained most of the visual stimuli (e.g. experimenter, windows); a hypersensitivity to these stimuli would explain a downcast gaze. This hypothesis is supported by the theory of Enhanced Perceptual Functioning (EPF) [14] which suggests that ASD children are overly sensitive to high frequency visual signals and proposes the use of an eccentric viewing strategy as a way to filter these signals [7]. Our findings in terms of exploration of the lateral field of view are consistent with this idea. Indeed, visual acuity diminishes gradually when directing the eyes sideways [15]. A strategy of eccentric viewing would therefore allow to reduce the sensory overload. This becomes important in the case of faces, which are a known element of atypical gaze behavior in autism (see [8] for a review). We are currently analyzing our data to evaluate to what extent lateral gaze is specifically used when the gaze is directed to the face of the experimenter. In contrast to [7], which reports only instances of extreme lateral gaze, our results seem to indicate the use of eccentric viewing as a continuous strategy across the whole field of view. Indeed, lateral glances correspond to when the gaze is directed to the extremes of the field of view. These occurrences do not suffice to explain the statistical differences we found in the two populations, as they only account for the outside parts of the field of view.

This study shows the advantages of using a tool such as the WearCam, to study the gaze of children in a natural interaction. The WearCam was designed for children as young as 1 year of age. It may hence be used to study the early development of vision and how it is affected by ASD. Moreover, it allows to study gaze across the whole field of view, giving us a greater access to how the world appears to children with ASD and, hopefully, a better insight on how they perceive what is around them.

## V. ACKNOWLEDGMENTS

We thank Martin Duvanel at EPFL and Myriam Laipe at CHUV for their work on the data analysis, and the parents and children who participated to the study. This work was funded by the Interactive Multimodal Information Management (IM2) project in the swiss National Centre for Competence in Research (NCCR) program.

## REFERENCES

[1] M. Rutter, "Incidence of autism spectrum disorders: Changes over time and their meaning*," *Acta Paediatrica*, vol. 94, no. 1, pp. 2–15, 2005.

[2] G. Bird, C. Catmur, G. Silani, C. Frith, and U. Frith, "Attention does not modulate neural responses to social stimuli in autism spectrum disorders," *NeuroImage*, vol. 31, no. 4, pp. 1614–1624, July 2006.

[3] J. Swettenham, S. Baron-Cohen, T. Charman, A. Cox, G. Baird, A. Drew, L. Rees, and S. Wheelwright, "The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants," *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 39, no. 05, pp. 747–753, 1998.

[4] A. Klin, W. Jones, R. Schultz, F. Volkmar, and D. Cohen, "Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism," *Archives of General Psychiatry*, vol. 59, pp. 809–16, 2002.

[5] E. R. Ritvo, D. Creel, A. S. Crandall, B. J. Freeman, C. Pingree, and R. Barr, "Retinal pathology in autistic children – a possible biological marker for a subtype?" *Journal of American Academy of Child Psychiatry*, vol. 25, p. 137, 1986.

[6] O. Bogdashina, *Sensory perceptual issues in autism and asperger syndrome*. Jessica Kingsley Publishers, July 2003.

[7] L. Mottron, S. Mineau, G. Martel, C. S.-C. Bernier, C. Berthiaume, M. Dawson, M. Lemay, S. Palardy, T. Charman, and J. Faubert, "Lateral glances toward moving stimuli among young children with autism: Early regulation of locally oriented perception?" *Development and Psychopathology*, vol. 19, pp. 23–36, 2007.

[8] D. R. Simmons, A. E. Robertson, L. S. McKay, E. Toal, P. McAleer, and F. E. Pollick, "Vision in autism spectrum disorders," *Vision Research*, vol. 49, pp. 2705–2739, 2009.

[9] Nadel and Butterworth, *Imitation in infancy*. Cambridge: Cambridge University Press, 1999.

[10] B. Noris, J.-B. Keller, and A. Billard, "Gaze direction detection for lay users in unconstrained environments," *Journal of Computer Vision and Image Understanding*, 2010.

[11] P. Mundy, C. Delgado, J. Block, M. Venezia, A. Hogan, and J. Seibert, *A manual for the abridged Early Social Communication Scales (ESCS)*, 2003.

[12] S. Sparrow, D. Balla, and D. Cicchetti, *Vineland Adaptive Behavior Scales, interview edition*. American Guidance Service, Circle Pines, MN, 1984.

[13] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

[14] L. Mottron, M. Dawson, I. Soulières, B. Hubert, and J. A. Burack, "Enhanced perceptual functioning in autism: An updated model, and eight principle of autistic perception." *Journal of Autism and Developmental Disorders*, vol. 36, pp. 27–43, 2006.

[15] J. A. M. Jennings and W. N. Charman, "Off-axis image quality in the human eye," *Vision Research*, vol. 21, pp. 445–455, 1980.