Unraveling the Developmental Dynamic of Visual Exploration of Social Interactions in Autism

Nada Kojovic[1,*], Martina Franchini[2], Holger Franz Sperdin[1], Corrado Sandini[1], Reem Kais Jan[3], Daniela Zöller[4], and Marie Schaer[1]

[1]Psychiatry Department, Faculty of Medicine, University of Geneva, 1211 Geneva, Switzerland
[2]Fondation Pôle Autisme, 1204 Geneva, Switzerland
[3]College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, 505055, United Arab Emirates
[4]Institute for Neuromodulation and Neurotechnology, Department of Neurosurgery and Neurotechnology, University Hospital Tuebingen and University of Tuebingen, 72076 Tuebingen, Germany
[*]Correspondence : Nada.Kojovic@unige.ch

Highlights

- A dynamic measure of divergence from normative gaze patterns was obtained while children with ASD watched animated cartoon
- The degree of divergence was negatively associated with adaptive and cognitive functioning
- During childhood viewing patterns in children with ASD were increasingly idiosyncratic
- Increased idiosyncrasy of gaze deployment in children with ASD characterized processing of sequences comprising social interaction

Summary

Atypical deployment of social gaze is present early on in toddlers with autism spectrum disorders (ASDs). Yet, studies characterizing the developmental dynamic behind it are scarce. Here we used a data-driven method to delineate the developmental change in visual exploration of social interaction over childhood years in autism. Longitudinal eye-tracking data were acquired as children with ASD and their typically developing (TD) peers freely explored complex animated scene. We found divergent moment-to moment gaze patterns in children with ASD compared to their TD peers. This divergence was particularly evident in sequences that displayed social interactions between characters, and even more so in the children with lower developmental and functional levels. Basic visual properties of the animated scene did not accounted for the enhanced divergence. Over childhood years, these differences dramatically increased to become more idiosyncratic. Taken together, these findings suggest that social attention should be targeted early in clinical treatments.

Keywords— autism spectrum disorders, eye tracking, social attention, data-driven, norm, development, social intensity, salience, divergence, idiosyncrasy

Introduction

Newborns orient to social cues from the first hours of life. They show privileged attention to faces (Simion et al., 2001), face-like stimuli (Goren, Sarty, and Wu, 1975; Johnson et al., 1991; Valenza et al., 1996), and orient preferentially to biological motion (Simion, Regolin, and Bulf, 2008). This automatic and preferential orientation to social cues early in life is highly adaptive as it provides grounds for the development of experience-dependent competences that are critical for the adequate functioning of an individual. Social visual engagement is one of the first means of exploration and interaction with the world, preceding and determining more advanced levels of social interaction and autonomy (Klin, Shultz, and Jones, 2015). Impairments in this elemental skill are one of the core characteristics of Autism Spectrum Disorders (ASD), a highly heterogeneous lifelong neurodevelopmental condition (American Psychiatric Association, 2013). Broad impairments in social communication and interaction along with repetitive behaviors and circumscribed interests have been suggested to lead to a spectrum of functional disabilities in ASD (Klin et al., 2007). In this regard, atypical social attention strategies may at least partially contribute to the emergence of the ASD phenotype. A large body of studies using eye-tracking have explored the atypicalities in attentional processes and their contribution to core symptoms in ASD (Chawarska and Shic, 2009; Falck-Ytter, Bölte, and Gredebäck, 2013; Klin et al., 2003). Recent meta-analyses concluded that, besides generally reduced social attention (Chita-Tegmark, 2016b), autism is also characterized by atypical attention deployment during exploration of social stimuli (Chita-Tegmark, 2016a). Indeed, aside from a generally diminished interest in social stimuli, when individuals with ASD do attend to social information they spend less time exploring key features, such as eyes while showing increased interest in less relevant cues such as bodies (Chita-Tegmark, 2016a,b). These atypicalities are present early on and exert tremendous impact on downstream developmental processes that critically depend on experience. The exact biological mechanisms that govern the emergence of these aberrant social attention patterns and their course of evolution are currently unknown.

In typical development, following the initial social preference, social attention deployment shows dynamic changes during infancy and early childhood. During their first year of life, infants progressively increase the time spent looking at faces compared to other elements of their environment (Frank, Vul, and Johnson, 2009). The increasing ability to attend to faces in complex environments has been related to developmental changes in visual attention (Frank, Amso, and Johnson, 2014). Indeed, during the first year of life, we observe the development of more endogenous, cortically controlled attention (Colombo, 2001), which allows more flexible and controlled displacement of gaze (Frank, Amso, and Johnson, 2014; Helo et al., 2016; Hendry et al., 2018; Hunnius and Geuze, 2004). Developmental improvement in attentional abilities thus promote engagement with social targets. Furthermore, the increase in capacity to attend to highly relevant social elements is followed by increased similarity in fixation targets between typically developing (TD) children (Frank, Amso, and Johnson, 2014). With increasing age, the TD children show more coherence in their visual behavior, as they increasingly focus on similar elements of the scene (Franchak et al., 2016; Frank, Vul, and Johnson, 2009; Shic et al., 2008). A trend toward progressively more coherent gaze patterns then continues into adulthood (Kirkorian, Anderson, and Keen, 2012; Rider et al., 2018). In other words, despite the impressive complexity of our social environment and despite the diversity of each individual's experiences, the social visual engagement takes a convergent path across the TD individuals, who are increasingly looking at the similar elements of the social environment. However, current understanding of the dynamic of this progressive tuning of gaze patterns is limited by the scarcity of the studies using longitudinal designs. Indeed, most studies used cross-sectional designs when inferring developmental patterns, which can be biased by interindividual differences.

In regards to autism, understanding the typical development of social visual exploration is of utmost importance, as the social difficulties associated with ASD are considered to result from the cascading effect of a reduced social interest during the child's development (Chevallier et al., 2012; Dawson et al., 1998; Dawson, Webb, and McPartland, 2005). Studies focusing on the developmental changes in visual exploration in autism are still rather scant, but point to altered maturational changes in orienting to social cues. Attention deployment begins to differ from the age of 2 months in babies who later develop autism suggesting that divergent trajectories of social visual exploration may start in the first months of life (Jones and Klin, 2013). A study by Shic et al. (2008) highlighted the absence of typical maturational change in face scanning strategies in children with ASD between 2 and 4 years of age. Longitudinal studies focusing on both typical and atypical development are thus crucially needed to highlight the underlying developmental mechanisms of atypical attention deployment in ASD. Longitudinal follow-up design would allow the identification of periods of critical changes in visual behavior that can be targeted by early interventions. In addition to the parsing of the developmental patterns, a comprehensive characterization of factors that influence visual behavior in social context is necessary to understanding the mechanisms of atypical attention deployment in autism.

Gaze deployment is mediated by numerous factors acting simultaneously and including both bottom-up and top-down processes. Bottom-up mechanisms direct attention to visually prominent elements as a function of their basic properties (such as orientation, intensity, color, motion) (Itti and Koch, 2000, 2001; Koch and Ullman, 1985) while top-down factors (Itti et al., 2001) are more endogenous in nature and depend on previous experience, motivation, specific task demands etc. (Yarbus, 1967). The complex interplay between these two groups of processes orchestrates our attention deployment during everyday tasks. We can hypothesize that the imbalance, such as enhanced focusing on bottom-up properties of visual content, may be at origin of atypical social attention in autism driving it away from conventional social targets. Indeed, it has been shown that in the context of naturalistic static scenes, individuals with ASD tend to focus more on

basic, pixel-level properties than on semantic categories, compared to TD individuals (Amso et al., 2014; Wang et al., 2015). However, less is known of the contribution of these basic properties to a real-time visual exploration of dynamic content as static stimuli only allow limited inference to the real world dynamic deployment of attention. Studies using dynamic social content are rare and point to somewhat contrasting results compared to the ones using static stimuli. For example, in the context of dynamic social content individuals with ASD tend to focus less on motion properties of the scene and more on intensity in comparison to TD children (Shic et al., 2007). However, there is currently no consensus in the literature on the relative predominance between bottom-up and top-down properties in generating aberrant visual exploration. These two processes were mostly analyzed separately and studies using ecological dynamic stimuli are scarce. Hence, another important element to consider is the content type as it dramatically influences the attentional processes that are summoned. For instance, non-social content is prone to elicit more heterogeneous patterns of exploration (Wang et al., 2018). On the other hand, social content of higher complexity induces more divergence in gaze deployment in TD (Wang et al., 2018) while giving rise to atypicalities in visual attention deployment in ASD (Chawarska, Macari, and Shic, 2012; Chita-Tegmark, 2016b).

Measures of gaze deployment (e.g. time spent on the face, or eyes), provided valuable insight into the specificity of social attention patterns in autism (Klin et al., 2002). These measures reflect the "macrostructure" (Guillon et al., 2014) of the gaze deployment by quantifying the overall time spent exploring predefined region of the scene. However, complementary to the "what" of gaze, the "when" of it is of equal importance as the demands in the real world come in an online fashion and require timely response. We attend to a only limited amount of elements from a breadth of possibilities and what finds the way to our perception will dramatically influence the meaning we attribute to the social situation. Recent studies have provided important advances in our understanding of the mechanisms that control what we select to gaze upon on a moment-to-moment basis (Constantino et al., 2017; Kennedy et al., 2017). Quite strikingly, while viewing social scenes, pairs of twins showed a high concordance not solely in the direction but also in timing of their gaze movements (Constantino et al., 2017; Kennedy et al., 2017). Thus, subtle variations in the visual exploration of social scenes are strongly influenced by genetic factors, that in turn favor the selection of important social information (Constantino et al., 2017). The process of continuous active selection of pertinent elements from the abundance of possibilities is critical for the interactive specialization of our brain (Johnson, 2001) and affects importantly how our internal world is shaped. Only few studies tackled the question of the moment-to-moment gaze deployment in ASD compared to TD. Indeed, while on this microstructural level TD individuals show coherence in fixation targets, fine-grained gaze dynamic in individuals with ASD is highly idiosyncratic and heterogeneous (Avni et al., 2019; Falck-Ytter and Hofsten, 2011; Nakano et al., 2010; Wang et al., 2018). Atypicalities in the fine-grained extraction of social information may have important consequences on learning opportunities and social functioning (Schultz, 2005). Overall, these findings urge for a better characterization of the underlying mechanisms and factors that contribute to coherence in visual patterns in typical development at different timescales, over months and years but also at the microstructural level (moment-to-moment) as a gateway for understanding the emergence of atypical gaze patterns in autism.

In the current study we opted for a comprehensive approach to characterize atypical visual exploration in children with ASD (1.8 - 6.9 years old) in comparison to their age-matched TD peers by considering both bottom-up and top-down processes. We first measured the divergence from normative gaze patterns in autism both on a microstructural level (moment-to-moment) as well as over larger temporal scales, measuring the developmental change during early childhood. We quantified the divergence between gaze patterns among the two groups of children while watching a cartoon depicting social interaction using a custom data-driven approach used in our previous studies (Jan et al., 2019; Kojovic et al., 2019; Sperdin et al., 2018). We further measured the developmental change in visual exploration in young children with ASD and their TD peers using the follow-up eye-tracking recordings collected yearly. Then we estimated the relative contribution of basic basic visual properties on the one hand, and intensity of social content on the other, to visual exploration of this dynamic social scene in both groups.

Results

Less divergent visual exploration was associated with better overall functioning in children with ASD

Gaze data from 59 males with ASD (3.74 ± 1.4 years) were recorded while children watched a 3-minute episode of the French cartoon *Trotro* (Lezoray, 2013). The cartoon depicts social interaction between three characters in a relatively slow pace. Our aim was to capture the subtlety of a moment-to-moment gaze deployment in ASD children compared to the TD group while watching this social scene. This was achieved by comparing the gaze allocation of each child with ASD to the normative gaze patterns obtained from 35 age-matched TD males (3.78 ± 1.3 years) who watched the same social scene. Normative gaze patterns (norm) were obtained by applying probability density estimation function (Botev, Grotowski, and Kroese, 2010) on gaze data from the TD group on each frame. Hence, for each child with ASD, we obtained a measure indicating the closeness to the norm that we denote *Proximity Index-PI,* (see Figure 1, Proximity Index Movie and Methods for detailed explanation). Lower PI values indicate a higher divergence from the norm for this complex social scene. In this manner we obtained a measure that indicated the divergence in gaze allocation compared to the group of age-matched typically developing children that, however, was not biased by author's theoretical priors

inherent to the approaches measuring time spent exploring predefined Areas of Interest (AOIs) (e.g. faces). Moreover, as it will be further detailed, this method allowed to flexibly redefine the normative gaze distribution by constraining the normative sample to a specific age range but additionally, to a specific group as well.

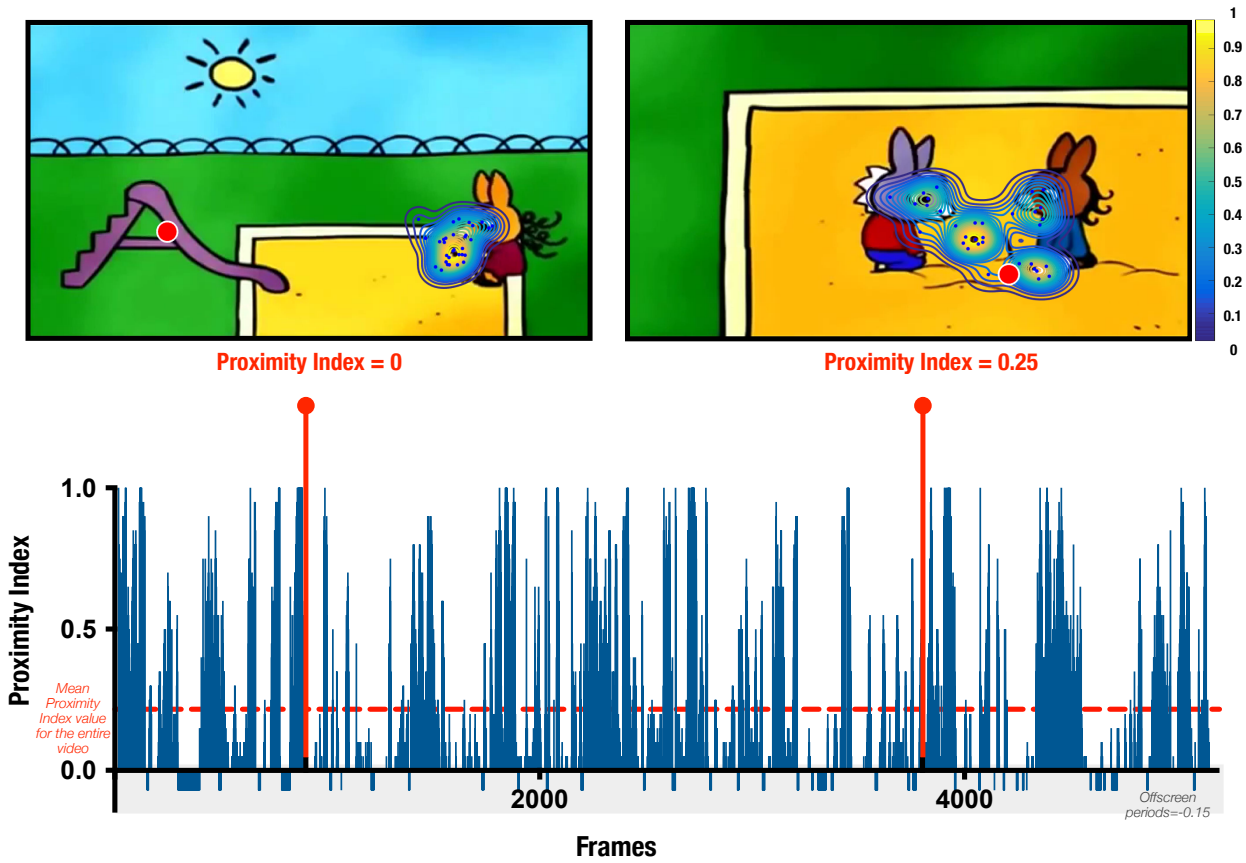

Figure 1: Proximity Index method illustration. Normative gaze data distribution ("norm") was created using gaze coordinates from 35 TD males (aged 3.79 ± 1.3 y.o.).
Upper row: two example frames with norm (delimited by contours) and gaze data from a 5 y.o. male with ASD (depicted as red circle). Hotter contour color indicates the area of higher density, meaning that a particular area was more appealing for a higher number of TD preschoolers for the given frame; the Proximity Index value for the 5 y.o. male with ASD for the frame on the left had a value of 0 and for the frame on the right a value of 0.25.
Lower row: Proximity Index values for the visual exploration of the 5 y.o. boy with ASD over the entire video with mean Proximity Index value indicated by the dashed red lines.

Individuals with ASD often present lower levels of adaptive functioning (Franchini et al., 2018; Hus Bal et al., 2015) and this despite cognitive potential (Klin et al., 2007). Understanding factors that contribute to better adaptive functioning in very young children is of utmost importance (Franchini et al., 2018) given the important predictive value of adaptive functioning on later quality of life. In our cohort, child adaptive behavior was assessed across communication, socialization, daily living skills, and motor domains using The Vineland Adaptive Behavior Scales, Second Edition, (Sparrow, Balla, and Cicchetti, 2005). Children with ASD showing visual exploration patterns that were less divergent compared to their TD peers had better adaptive functioning profile as assessed by VABS-II ($r = 0.399$, $p < 0.001$) (see Figure 2). This correlation was driven by positive correlations in subdomains of communication ($r = 0.263$, $p = 0.044$), socialization ($r = 0.354$, $p = 0.006$) and autonomy in daily life tasks ($r = 0.301$, $p = 0.020$) (see Figure S1). The correlation in the communication subdomain was further driven by the association between divergence in gaze and language understanding ($r = 0.299$, $p = 0.021$) and expressive language skills ($r = 0.435$, $p < 0.001$). There was no significant association between PI values and written language skills (third subdomain of the communication domain).

In the context of the developmental assessment, more TD-like visual exploration characterized children with ASD who showed better communication ($r = 0.398$, $p = 0.001$) and motor skills as assessed by PEP-3 (Schopler, 2005) ($r = 0.294$, $p = 0.015$). As depicted on Figure S2, all three scales that composed the communication composite domain

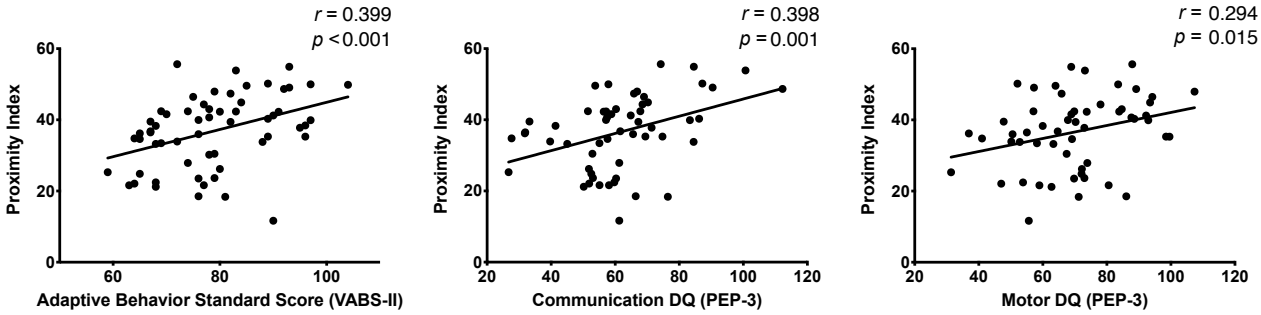

Figure 2: Proximity Index and its relation to behavioral phenotype in children with ASD. Correlation between Proximity Index-PI (based on normative gaze data from 35 TD children) and measures of Adaptive functioning (VABS-II) and Communication and Motor skills developmental quotients (PEP-3) for the group of 59 preschoolers with ASD age-matched to the normative sample.

of PEP-3 were positively related with PI (Verbal and Preverbal Cognition: $r = 0.307$, $p = 0.023$, Receptive Language: $r = 0.381$, $p = 0.004$ and Expressive Language: $r = 0.304$, $p = 0.024$). In the motor domain, less divergence in gaze was related to better imitation skills ($r = 0.350$, $p = 0.004$) and fine motor skills ($r = 0.286$, $p = 0.034$. There was no association between the PI and global motor skills ($p = 0.15$).
Finally, no significant correlation between the PI and the severity of autistic symptoms was found (r = -0.17, $p = 0.187$).

Divergent developmental trajectories of visual exploration in early childhood in children with ASD

Our next aim was to discern the developmental pathway of the visual exploration in this complex social scene for both groups of children. Previous studies using cross-sectional designs have demonstrated important changes in the way children attend to social stimuli depending on their age (Frank, Vul, and Saxe, 2012; Helo et al., 2014). As our initial sample spanned a relatively large interval (1.8 - 6.9 years), we wanted to obtain a more fine-grained insight into the developmental dynamic of visual exploration during the given period. To that end, when study-specific inclusion criteria were satisfied, we included longitudinal data from our participants who had a one year and/or a two years follow-up visit (see Methods). With the available 106 recordings for the ASD group and 78 for the TD group we applied a sliding window approach (Sandini et al., 2018) (see Methods). Precisely, our goal was to discern critical periods of change in visual exploration of complex social scenes in ASD compared to TD group. We opted for a sliding window approach considering it's flexibility to derive a continuous trajectory of visual exploration and thereby capture such non-linear critical periods. The sliding window approach yielded a total of 40 age-matched partially overlapping windows for both groups covering the age range between 2.2 - 4.5 years (mean age of the window) (please refer to Figure S3 for the illustration of the sliding window method).

We then estimated gaze dispersion on a group level across all 40 windows in two groups using two complementary measures. The first measure, Median Contour Surface (MCS), depicts the surface of the median level contour (isoline) projected on the estimated gaze density function. The second, the Convergence Index (CI), indicates the "steepness" of the distribution (see Methods for a detailed description of these two measures). To measure the trajectory of gaze dispersion, models of increasing order (constant to cubic) were fit on each of the two measures of dispersion. Bayesian Information Criteria (BIC) was used to select the best fitting model (Mutlu et al., 2013). Both measures of dispersion, MCS and CI, indicated an increasingly discordant pattern of visual exploration between groups during early childhood years (see Figure 3). While the TD children showed more convergent visual exploration pattern as they got older, as revealed by progressively smaller values of MCS (narrowing of focus) and higher values of CI (steeper distribution of density function), the opposite pattern characterized gaze deployment in children with ASD. From the age of 2 years up to the age of 4.5 years this group showed a progressively discordant pattern of visual exploration ($p < 0.0001$, see Figure 3, see replay illustrating dispersion change across all 40 windows for the frame showing two characters interacting : Sliding Window Movie).

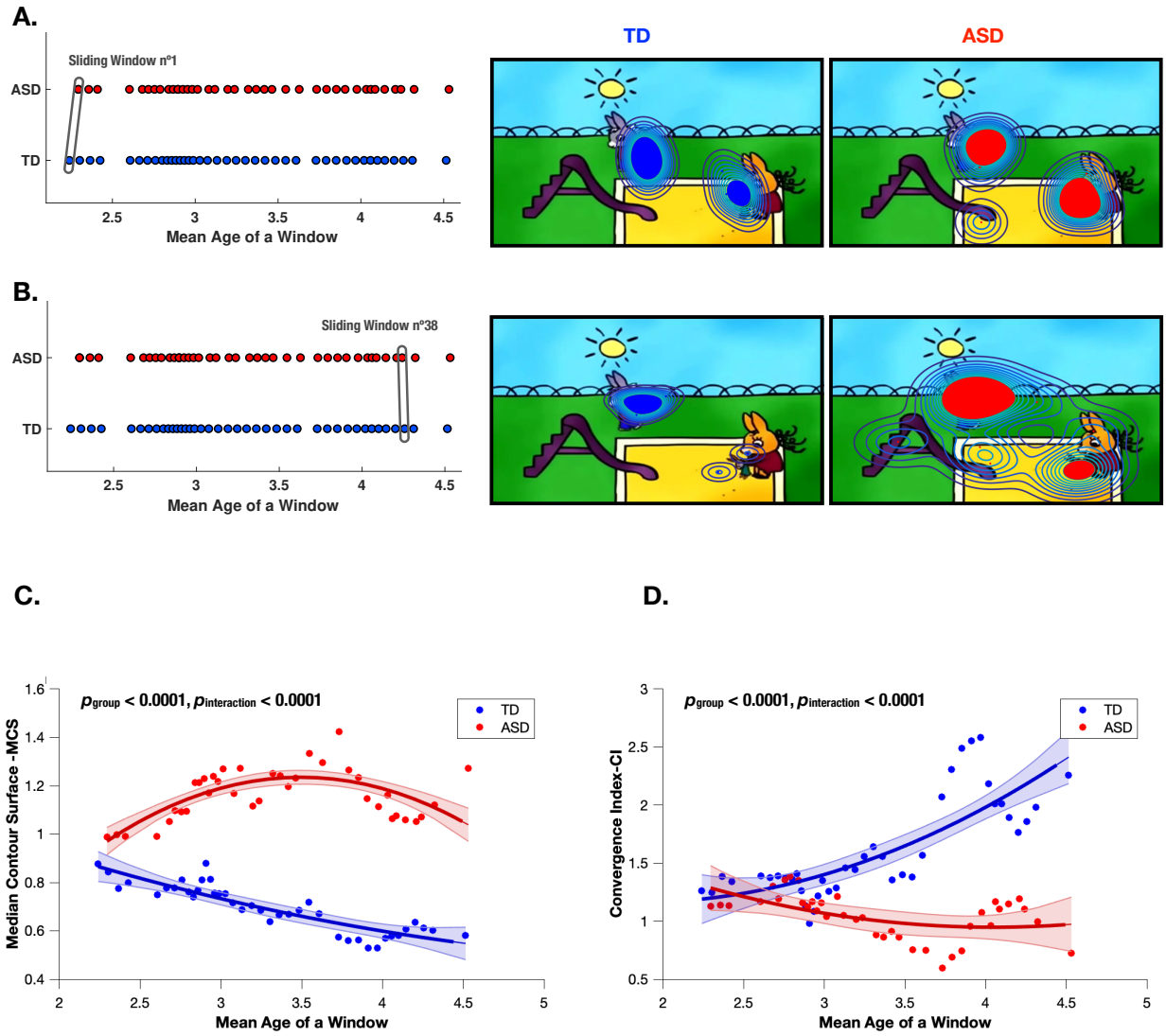

Figure 3: Characterization of the evolution of visual exploration patterns in young children with ASD and the TD group using a sliding window approach.
Panels A & B illustrate gaze dispersion in two groups for the sliding window n°1 and n°38 (mean age of windows 2.2 and 4.3 years, respectively); each circle represents a window encompassing 16 recordings;
Panel C shows the comparison of the gaze dispersion between two groups using Median contour surface (MCS) measure. MCS was calculated across 40 sliding windows spanning 2.2-4.5 years of age on average (here again every circle represents a window encompassing 16 recordings).
Panel D depicts comparison of the gaze dispersion between two groups using our second measure of dispersion, Convergence Index (CI). As can be seen on panel C, MCS values diminished in TD group with advancing age while the opposite pattern is observed in the ASD group. This would suggest a progressively more dispersed gaze behavior in the ASD group during childhood years. Similarly, as depicted on panel D, CI reflects the "steepness" of the distribution and tended to increase in the TD group with age reflecting increasingly fine tuned visual exploration in this group over childhood years.

One of the characteristics of the method we employed to estimate gaze density distribution is its ability to dynamically capture the coexistence of several attentional foci for the given group. By focus we denote a hotspot in a gaze distribution that many children were attending on a given frame (see Figure 1 for an example of monofocal (left) and multifocal (right) gaze distribution in a group of 35 TD children on a frame depicting social interaction episode). Thus, our next aim was to shed more light on the evolution in gaze distribution in both groups by comparing additional gaze dispersion measures (see Figure S4). We analyzed the evolution of the overall number of separate attentional foci at three

different levels of the gaze distribution (delimited by lowest, median and highest level isolines) as well as the evolution of the respective surface covered by the given focus over the childhood years in both groups (for a detailed explanation please refer to Supplementary section). Over the childhood years gaze patterns in the ASD group were indicative of a progressive loss of the group coherence. Thus with increasing age the overall surface covered by the gaze distribution estimation in this group increased and subsequently reached a plateau while the overall number of higher level foci (indicating convergence) decreased from the age of 3 years old.

Next we wanted to validate the findings on dispersion by using a longitudinal analysis that is less prone to confounds such as biased sampling. To this end we conducted a dispersion analysis on our smaller one-year follow-up longitudinal sub-sample. This sub-sample included 34 males with ASD (3.5 ± 1.2 years old) and 24 age-matched TD peers (3.3 ± 1.3 years old) for whom we had two recordings obtained a year apart. We compared both groups at baseline as well as at follow-up with regards to the two dispersion measures (Median contour surface and Convergence Index) using Wilcoxon matched-pairs signed rank test. For all inter and intra-group comparisons, Wilcoxon test returned significant, with $p < 0.0001$. The two groups were significantly different at baseline with regards to both dispersion measures, effect sizes are given for MCS and CI respectively ($r_{mcs} = 0.16$, $r_{ci} = 0.07$ but the group gap became more pronounced a year later as indicated by higher effect sizes ($r = Z/\sqrt{N}$, (Rosenthal, 1991)) of group difference $r_{mcs} = 0.38$, $r_{ci} = 0.32$) (see Figure S5). Here again, the TD group showed more fine tuned gaze patterns over a period of one year compared to the ASD group ($r_{mcs} = 0.26$, $r_{ci} = 0.25$). Although the effect size was fairly small ($r_{mcs} = 0.04$, $r_{ci} = 0.05$) when compared to baseline, the ASD group showed a more dispersed gaze pattern a year later (see Figure S5).

Finally, we analyzed the within-subject change in gaze exploration over the period of one year using the same one-year follow-up longitudinal sub-sample. Here, gaze data from the TD group at both time points (baseline and one-year follow-up) were used to define norms against which we then compared gaze data of children with ASD both at baseline (T1) and at follow-up (T2). As can be seen on Figure 4 Panel B, gaze patterns in males with ASD at one year follow-up (ASD T2) aged 4.6 ± 0.5 were more divergent from the age-matched norm at follow-up (TD T2) $M_{TDT2}^{ASDT2} = 0.386$ (± 0.087) than from the age-matched norm at baseline (TD T1)($M_{TDT1}^{ASDT1} = 0.427$ (± 0.100), $t(33) = 13.63$, $p < 0.0001$, $r = 0.95$). As shown on Figure 4, Panel C, when gaze deployment at T2 was compared to the now younger norm (TD T1), no significant difference was evidenced with regards to Proximity index values children obtained a year earlier ($p > 0.05$) suggesting aberrant maturation processes of gaze deployment in ASD group.

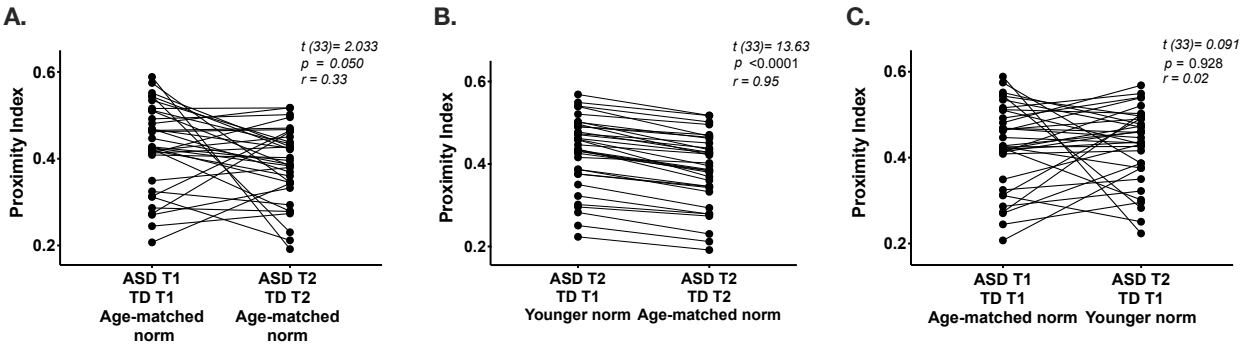

Figure 4: Measure of a within-subject change in PI over a period of one year in children with ASD. A. Comparison of Proximity Index values between two time points (norm is composed of gaze data of age-matched typically developing children); B. Values of Proximity index for the group of ASD children at follow-up were obtained using a norm at follow-up (Norm TD T2) and a norm at baseline (Norm TD T1); C. Values of Proximity index for the group of children with ASD at baseline obtained using norm at baseline (Norm TD T1) were compared to Proximity Index values from children with ASD obtained using a year younger norm (TD T1) as reference.

More ambient and less focal fixations in children with ASD compared to the TD group

Next we wanted to complement our analysis using standard measures of visual behavior. In our cross-sectional sample of 59 males with ASD (3.74 ± 1.4 years) and 35 TD children (3.78 ± 1.3 years), we did not find any significant difference between groups with regards to overall number of fixations, saccades, median saccade duration, nor saccade amplitude for the duration of the cartoon ($p > 0.05$). However, median fixation duration was higher in TD children compared to the ASD group ($t(92) = 2.93$, $p = 0.004$) suggesting a more focused attentional style in the TD group. To characterize the predominant attention exploration mode while watching the cartoon, we defined two types of fixations, based on their duration and the length of the saccade preceding them. Thus using thresholds as in (Unema et al., 2005), a fixation was considered as "focal" if longer than 180ms and preceded by a saccade of an amplitude smaller than 5° of visual angle. Shorter fixations (< 180ms) and preceded by a longer saccade (> 5°) were classified as "ambient". We then obtained the proportion of these two fixation types normalized for the overall fixation number. In the ASD group, we

observed significantly more ambient fixations ($U = 500$, $p < 0.0001$) compared to the TD group. The TD group showed more focal fixations ($U = 497$, $p < 0.0001$) in comparison to the ASD group. In both groups, focal fixations were more frequent than ambient ($p < 0.0001$) (see Figure S6 A1). Higher presence of focal fixations was related to higher values of Proximity Index ($r = 0.483$, $p < 0.001$) while the opposite relationship was evidenced between Proximity index and proportion of ambient fixations (r = -0.395, $p = 0.002$)(see Figure S6 Panels A2&3).

To measure the developmental changes in these two attentional modes, we used our previously described unstructured longitudinal sample including 106 recordings for the ASD group and 78 for the TD group. Random slope mixed model analysis (Mancini et al., 2019; Mutlu et al., 2013) was employed for focal and ambient fixations, respectively (see Figure S6 Panels B1 &$B2$). Over the childhood years, the TD group used consistently more focal fixations ($p_{group} < 0.001$). In contrast, the ASD group used more ambient fixations than the TD group ($p_{group} = 0.004$).

Basic visual properties of the animated scene were not at the origin of the enhanced divergence

We next estimated the relative contribution of basic visual properties of the scene to visual exploration in both groups. Previous studies in adults with ASD have shown that these basic properties play an important role in directing gaze in ASD individuals while viewing naturalistic images (Amso et al., 2014; Wang et al., 2015). Less is known about the contribution of these properties to gaze allocation in animated scenes during early childhood as well as the developmental dynamics of this contribution. Therefore, we extracted the values of five salience features (intensity, orientation, color, flicker, motion) for each frame of the video using two variants of biologically inspired salience model, namely Graph Based Visual Saliency (GBVS) (Harel, Koch, and Perona, 2007) and the standard Itti and Koch model (I&K model) (Itti and Koch, 2001; Itti, Koch, and Niebur, 1998) as explained in details in the Methods section. Firstly, we calculated salience measures for our cross-sectional sample with 59 males with ASD and age-matched 35 TD males on a group level for each frame. For each channel (intensity, orientation, color, flicker and motion) as well as the full model (linear combination of all five channels), we calculated the area under a Receiver Operating Characteristic curve (ROC) (Green and Swets, 1966). GBVS model preformed slightly better than the I&K model in the prediction of the gaze so for the sake of simplicity we will report only the results obtained using this model. Contrarily to our hypothesis, for all channels taken individually as well as for the full model, the salience model better predicted gaze allocation in the TD group compared to the ASD group (Wilcoxon t-test returned with the value of p < 0.0001). The effect sizes of this difference were most pronounced for the orientation channel $r = 0.612$, followed by the full model $r = 0.611$, intensity $r = 0.609$, color $r = 0.609$, motion $r = 0.582$ and lastly flicker channel $r = 0.575$.

To gain insight into the individual patterns of gaze allocation with regards to salience as well as the developmental trajectory, we obtained values of the area under a ROC curve for each participant individually on each frame. This was done on our unstructured longitudinal sample including 106 recordings from children with ASD and 78 recording from TD children. Taking into account the variable number of the recordings available for each subject (1-3), a longitudinal mixed model regression (Mancini et al., 2019; Mutlu et al., 2013) was conducted in order to delineate the trajectory of the influence that basic visual properties of the visual scene exert in gaze allocation for the two groups. False Discover Rate (FDR) was used to correct for multiple comparisons (Genovese, Lazar, and Nichols, 2002). As shown on Figure 5, gaze allocation was better predicted by salience values in TD children as compared to the children with ASD in all used channels and in a full model. Quite strikingly during the childhood year, the TD group relied consistently more on salience properties for all channels taken at once ($p_{group} = 0.042$), as well as in the intensity ($p_{group} = 0.004$), orientation ($p_{group} = 0.012$), motion ($p_{group} = 0.011$), flicker ($p_{group} = 0.021$) and lastly color channel ($p_{group} = 0.047$) compared to children with ASD. Group differences appeared stable across childhood with no significant interaction effect with age.

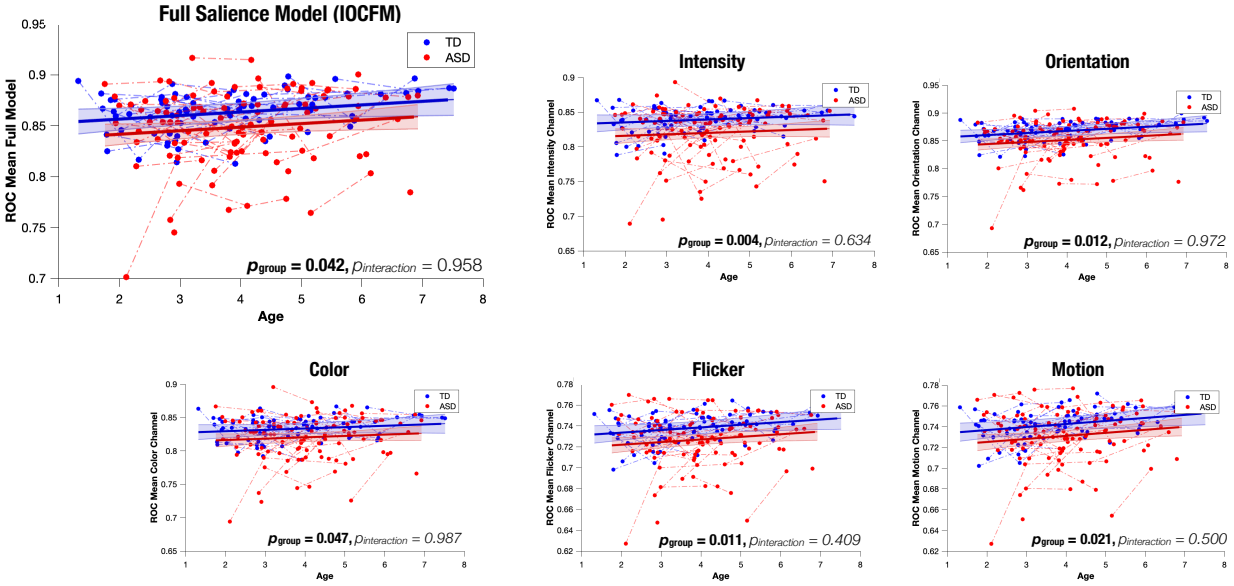

Figure 5: Mixed-model regression analysis over the mean ROC curve for 106 recordings from preschoolers with ASD (4.03 ± 1.25) and 78 recordings from TD children (3.81 ± 1.47) for full GBVS salience model (all channels combined) and 5 individual channels (Harel, Koch, and Perona, 2007). The group-level trajectory (solid line) is derived from mixed-effect modeling (Mancini et al., 2019; Mutlu et al., 2013). The colored bands indicate the 95% confidence interval around the estimated group-level trajectory. Significant findings are highlighted in bold, no interaction between age and salience features was detected. TD group gaze allocation was consistently better predicted by salience features as compared to the ASD group.

As previously, to precisely quantify the within-subject change in salience contribution to the allocation of gaze, we constrained our analysis to the one-year follow-up longitudinal sub-sample of 34 children with ASD and 24 TD children (see Figure S7). Concordant with our previous findings here the Full model gained in efficiency and predicted significantly better gaze allocation in TD children a year after the initial visit, that it did in children with ASD ($p = 0.040$). Year after the initial assessment all individual channel's contribution to the prediction of the gaze allocation in TD group were significantly improved in comparison to the ASD group (intensity $p = 0.008$, orientation $p = 0.022$, motion $p = 0.031$, flicker $p = 0.036$ and finally color channel $p = 0.047$).

Intensity of social content was associated with divergence in visual exploration in ASD group

Taking into account previous findings of enhanced difficulties in processing more complex social information (Chita-Tegmark, 2016b; Frank, Vul, and Saxe, 2012; Parish-Morris et al., 2019) in individuals with ASD we tested how intensity of social content and visual complexity influenced visual exploration of the given social scene. As detailed in Methods section, social complexity was defined as the total number of characters for a given frame and ranged from 0 to 3 for this video. Frames with no characters represented a substantial minority (0.02% of total video duration) and were excluded from further analysis. In our cross-sectional sample, we found that PI values were significantly higher for frames containing one character as compared to those with two or more ($M_{1char} = 0.395 \pm 0.103$, $M_{2/3chars} = 0.3387 \pm 0.105$, $t(58) = 11.29$, $p < 0.0001$, $r = 0.83$). No significant difference in Proximity Index was found between frames containing 2 vs 3 characters so these were merged (see Figure 6A).

Additionally, we analyzed the influence of the overall visual complexity of the scene to this divergent visual exploration in the ASD group. The total length of edges defining details on the images was employed as a proxy for visual complexity (see Methods for more details). After standardization, social intensity was entered as regressor in a General Linear Model (GLM) analysis conducted in Matlab (http://www.mathworks.com) with Proximity Index as dependent variable. Social complexity emerged as a significant predictor of Proximity Index values ($\beta = 0.237$, $p < 0.0001$). Entering visual complexity values in the model did not improve the prediction of Proximity Index values ($\beta = 0.002$, $p = 0.908$). While watching this animated social scene the overall richness in details did not influence the divergence from the normative gaze patterns in the ASD group. However, more intensity in social content such as moments of interaction between two or three characters was related to greater divergence in gaze deployment in the ASD group, suggesting aberrant involvement of top-down attentional processes.

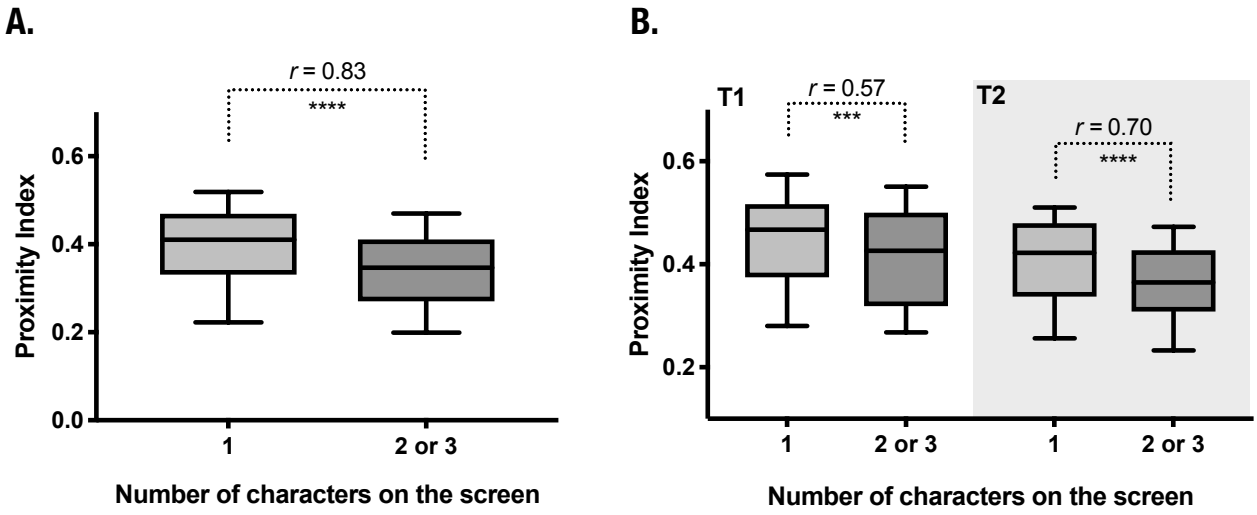

Figure 6: Contribution of social intensity to gaze divergence in children with ASD. A. Proximity values comparison in the cross-sectionnal sample of 59 children with ASD and 35 TD children with regards to two levels of social intensity: frames depicting one character alone and frames depicting interaction (these included two or three characters); B. Comparison of mean Proximity values in a one-year follow-up longitudinal sample of 34 children with ASD and 24 TD children with regards to two given levels of social intensity (on the left comparison at the baseline and on the right a comparison of the proximity values obtained a year after the initial visit).

Lastly, we characterized the within-subject change with regards to social content intensity. As shown on Figure 6B results from the one-year follow-up longitudinal sample corroborate the findings of negative impact of social intensity on the Proximity Index, both at baseline and a year later noticing an important increase in effect size a year later (at baseline: $M_{1char} = 0.442 \pm 0.103$, $M_{2/3chars} = 0.411 \pm 0.105$, $t(34) = 3.399$, $p = 0.0003$, $r = 0.57$; a year later: $M_{1char} = 0.405 \pm 0.094$, $M_{2/3chars} = 0.361 \pm 0.090$, $t(33) = 5.576$, $p < 0.0001$, $r = 0.70$). More intensity in social information was more strongly related to divergence in visual exploration in ASD a year later compared to baseline. This increasing difficulty in processing socially intense information was indicative of the existence of aberrant maturation process.

Discussion

In the present study, we used a data-driven method to quantify differences in spatio-temporal gaze patterns between children with ASD and their TD peers while watching an animated movie. Children with ASD who showed less moment-to-moment divergence in the exploration of a 3-minute cartoon compared to normative gaze distribution of age-matched TD children, had better adaptive functioning and better communication and motor skills. On a larger temporal scale, across childhood years, the TD children showed a progressive tuning in the focus of their attention, reflected by a narrowing of the group focus while the ASD group showed increasingly dispersed gaze patterns over the same period. Strikingly, in children with ASD who were followed longitudinally over the period of one year, their visual exploration diverged from the one seen in the TD group and was closer to the same normative group as at baseline suggesting impaired maturation processes at stake. Contrarily to our expectation, visual exploration in the group of children with ASD was not better predicted by low level salience of the visual scene compared to their TD peers. These low level salience feature gained more influence in the gaze deployment of the TD children with age. However, this was not observed in the ASD group. Lastly, the intensity of social content had an important impact on divergence from the TD gaze patterns; children with ASD showed more divergent deployment of attention on scene sequences with more than one character suggesting difficulties in processing social cues in the context of social interaction.

Our results corroborate and extend findings of a growing body of studies that have explored microstructural gaze dynamics in autism (Avni et al., 2019; Falck-Ytter et al., 2013; Nakano et al., 2010; Wang et al., 2018) and have demonstrated divergent moment-to-moment gaze deployment in children with ASD compared to their age-matched TD peers. Any slight but systematic divergence in gaze deployment can have tremendous influence on the experience dependent brain specialization (Johnson, 2001; Klin et al., 2009). These subtle but relevant patterns might not be detected by methods focusing on macrostructural gaze structure measuring overall attention allocation on distinct visual features (e.g. faces, eyes, etc.) based on predefined Areas of Interest (AOI). Here, we extend the existing findings by using a different data-driven methodology and by including the developmental aspect of the spatiotemporal gaze deployment in autism and typical development. In our study, to define the normative gaze behavior we opted for an approach that accounts for multiple socially relevant targets simultaneously present on the screen (e.g. two characters interacting) thus providing more subtle estimation of the group gaze deployment (see Figure S4 B1). Furthermore, in this study we used a cartoon, which while not being a completely ecological presentation of social interactions, has the advantage of being very appealing to young children. Previous research (Riby and Hancock, 2009) has shown that children with ASD attend more to dynamic cartoon stimuli representing social interaction compared to when shown natural movies of people interacting. Despite animated movies being a simplified version of social interaction with reduced social complexity, the movie we analyzed provided us with ample insight on the atypicalities of gaze behavior in children with ASD.

The sliding window approach yielded a fine grained-measure of change in gaze deployment in both groups of children during early childhood. With advancing age, TD children showed increasingly coherent gaze patterns, corroborating previous findings of increased consistency in TD gaze behavior over time (Franchak et al., 2016; Frank, Vul, and Johnson, 2009; Kirkorian, Anderson, and Keen, 2012; Rider et al., 2018; Shic et al., 2007). On the other hand, children with ASD showed increasingly heterogeneous patterns during the same period. A similar contrasting pattern with gaze in TD individuals getting more stereotyped from childhood to adulthood and gaze in ASD groups showing more variability was brought forward in a study by Nakano et al. (2010). While this study used a cross-sectional design to study developmental change in a group of children and adults, to our knowledge our study is the first to extend the findings on both TD and those with ASD using longitudinal design and focusing on a moment-to-moment gaze deployment. Over the period of one year we noted a striking change in gaze patterns in TD preschoolers reflecting consistency and a tendency to look at similar elements of the social scene at the same time. This higher consistency in gaze in the TD group with increasing age was put in relation with more systematic involvement of top-down processes (Franchak et al., 2016; Helo et al., 2017; Kirkorian, Anderson, and Keen, 2012). During typical development through the phylogenetically (Rosa Salva et al., 2011) favored mechanism of preferential orientation to social stimuli, children show increasing experience with and subsequently increasing understanding of social cues setting them on the typical path of social development (Jones and Klin, 2013; Klin et al., 2009). On the other hand, strikingly divergent pattern in children with ASD might be seen as a product of accumulation of atypical experiences triggered by social attention deployment diverging early on in their development (Jones and Klin, 2013). Behaviorally, in children with ASD during the preschool years we observe the emergence of circumscribed interests alongside with the tendency of more rigid patterns of behaviors (insistence on sameness) (Richler et al., 2010). These emerging patterns of interests might contribute to the divergence in gaze by attracting attention to elements related to the circumscribed interests (Sasson et al., 2008, 2011) thus amplifying the derailment from the normative social engagement path (Klin, Shultz, and Jones, 2015). We could therefore expect that these specific patterns of interest could indeed yield a divergent (inter-group) but nevertheless consistent (intra-group) path of visual exploration in this population. Thus, on a group level we would expect to observe a certain level of the intra-group consistency centered around circumscribed interest-related targets. Ultimately, interest are, indeed, specific in nature and might not necessary lead to a group-level coherence, we could, however, expect a discernible amount of a within-subject stability in gaze patterns over shorter time scales. The results of the present study highlighted the loss of the group cohesion with age in gaze in children with ASD on a group-level which corroborates emerging findings of marked gaze inconsistency across individuals with ASD (Nakano et al., 2010; Wang et al., 2018) and is evidenced within-individuals when the same video was seen twice (Avni et al., 2019). In other words, our results do not corroborate

the hypothesis of competing attention to non-social stimuli being clearly at the origin of gaze atypicalities in ASD but argue more for the presence of growing idiosyncrasy in the selection and processing of information in social context in this population. An increased idiosyncrasy on the neural level while watching dynamic social scenes has been put forward by a number of studies (Bolton et al., 2020; Bolton et al., 2018; Byrge et al., 2015; Hasson et al., 2009) and was related with lower scene understanding (Byrge et al., 2015) and higher presence of autistic symptoms (Bolton et al., 2020). The mechanisms of efficient selection of relevant social elements are genetically controlled (Constantino et al., 2017; Kennedy et al., 2017) and a disturbance we observe in ASD is most likely a downstream product of the gene-environment correlation (Klin, Shultz, and Jones, 2015). According to this view, the initial vulnerability (Constantino et al., 2017; Jones and Klin, 2013) characterizing autism would lead to a life-time of atypical experiences with the social world which in turn could result in atypical brain specialization and more idiosyncratic behavioral patterns.

In the present study we showed that the level of divergence in gaze exploration of this 3-minute video was correlated with developmental level in children with ASD as well as their overall level of autonomy in various domains of everyday life. This finding stresses the importance of studying the subtlety of gaze deployment with respect to its downstream contribution to more divergent global patterns of development (Jones and Klin, 2013; Klin, Shultz, and Jones, 2015; Schultz, 2005; Young et al., 2009). With regards to the exploration style while watching the cartoon, compared to their TD peers, children with ASD presented more ambient, exploratory fixations, indicative of a rapid acquisition of low-frequency information (Eisenberg and Zacks, 2016). On the other hand, they showed significantly less focal fixations that are known to operate with more fine grained high-frequency information. This suggests that children with ASD spent more time than the TD group in an ambient mode trying to grasp the global scene configuration (Ito et al., 2017) and less in a detail-sensitive focused mode. These two modes of exploration are supported by distinct and yet functionally related systems of dorsal attention (ambient mode related processing of spatial relations) and ventral attention (dealing with behaviorally salient object representation through the involvement of focused mode) (Helo et al., 2014). Our finding of differential recruitment of these two modes during viewing of social stimuli might suggest different recruitment of these two attentional networks during processing of this complex social scene. In our previous work on a smaller sample for which we also acquired EEG recording during the time that children watched the Trotro cartoon, we found that the divergence in gaze deployment was related to the vast abnormalities in neural activation, including reduced activation of frontal and cingulate regions and increased activation of inferior parietal, temporal, and cerebellar regions (Jan et al., 2019). In a similar EEG-eye-tracking study using videos involving biological motion (children doing yoga in nature) (Sperdin et al., 2018), we found increased driving from regions such as the median cingulate cortex and the paracentral lobule in the toddlers and preschoolers with ASD who had a more similar visual exploration pattern to their TD peers (higher PI). Thus, the children who showed less divergence from normative gaze patterns (TD-like viewing patterns) engaged more importantly median cingulate cortex and the paracentral regions suggesting potential compensatory strategies to account for the divergent brain development over time. Longitudinal studies combining eye-tracking and neuroimaging techniques are necessary to confirm the hypothesis of compensatory hyperactivation.

In an effort to parse the growing idiosyncrasy in gaze deployment evidenced in our ASD group across childhood years, we measured the contribution of basic visual properties of the scene to the gaze deployment in this group as compared to the TD group. We found that the basic visual properties played a less important role in directing gaze in our group of young children with ASD as compared to their TD peers. This was observed across all separate channels, namely, intensity, orientation, color, motion and flicker as well as the full salience mode with all channels combined. Surprisingly, for both groups static channels contributed more to the gaze allocation than the two dynamic ones (motion and flicker). This is possibly due to the nature of the movie where motion is generated in a non continuous way, with 3 to 4 identical frames sometimes occurring before a change which is in marked contrast to how the movement is perceived in natural environments. Previous research has shown that bottom-up features are responsible for directing attention in very young infants but from 9 months of age top-down processes take the predominance in directing the gaze (Frank, Vul, and Johnson, 2009). Less is known about the relative contribution of these processes while watching complex dynamic stimuli over a developmental span. Using a cross-sectional sample of TD children and adults Rider et al. (2018) showed that gaze deployment in group of children and adults was better predicted by the presence of a face in the scene (summoning top-down processing mechanisms) than by low level visual properties of the scene. However, two salience models they used (I&K and GBVS, same as those used in our study) better predicted gaze data in adults than they did in children suggesting that these dynamic salience models might be more adapted to the mature visual system. Indeed our sample is relatively young and it is possible that the lesser success of the salience models to predict gaze allocation in ASD children might be influenced by the visual and motor abnormalities characterizing this group (Farber and Beteleva, 2005; Rider et al., 2018). Effectively, in our longitudinal one-year follow-up sample we noticed that the mean ROC values across all five channels and the full model showed a tendency to increase in one year's time in TD group suggesting that the models we used better behave in the context of more mature visual behavior. Nevertheless, both salience models were indeed conceived to match the mature adult visual system and might not deal as well as with visual systems at earlier stages of development.

Contrary to the bottom-up properties of the given scene, social intensity was an important element in governing the gaze divergence in children with ASD. The finding of more divergent pattern in frames comprising the interaction between characters corroborates previous findings of atypical face (Hanley et al., 2013) and dynamic social stimuli (Speer et al.,

2007) processing particularly in context of interaction (Parish-Morris et al., 2019). The enhanced divergence in moments including interaction was not better explained by the overall perceptual demands of the scene as the visual complexity of the given scene did not contribute importantly to the gaze allocation. Processing of social interaction depends strongly on the top-down inputs, as the choice of what is to be attended relies on prior expectations, attributed meaning and globally language and scene understanding. Over the period of one year children with ASD showed heightened divergence in their manner to attend to social interaction compared to baseline.

The finding of progressive divergence in gaze patterns in children with ASD during the childhood years urges for an early detection and early intensive intervention to prevent further derailment from typical social engagement path (Dawson et al., 2010). The present study is one of the first to tackle microstructural atypicalities in gaze deployment in young children with ASD taking into account developmental change. This work stresses the need to better characterize the link between behavioral phenotypes and the underlying neurobiological substrates in order to adapt early intervention strategies to the neurodevelopmental mechanisms involved.

Methods

Participants

Cross-sectional sample

Fifty-nine children with autism (3.74 ± 1.4 years) and 35 age-matched typically developing children (3.78 ± 1.3 years) participated in the study. Due to the fewer number of eligible females with ASD our study included only males. All children with ASD reached the cut-off for ASD on Autism Diagnostic Observation Schedule-Generic ADOS-G), (Lord et al., 2000) or Autism Diagnostic Observation Schedule-2nd edition (ADOS-2) (Lord et al., 2012). For subjects who underwent the ADOS-G assessment the scores were recoded according to the revised ADOS algorithm (Gotham et al., 2007) to ensure comparability with ADOS-2. Typically developing (TD) children were screened for the presence of any known neurological or psychiatric illness in any first-degree relative and the child itself. The data for the current study were acquired as a part of a larger longitudinal study of early development in autism, based in Geneva. Detailed information about cohort recruitment was given elsewhere (Franchini et al., 2017, 2018; Kojovic et al., 2019). This study protocol was approved by the Ethics Committee of the Faculty of Medicine of the Geneva University, Switzerland. All families gave written informed consent to participate.

Unstructured longitudinal sample

As participants to our study are followed longitudinally their repeated visits were included when satisfying the inclusion criteria (later detailed in the Method details section). This yielded a total of 106 recordings for the ASD group and 78 for the TD group (all recordings were collected a year apart; 38 children with ASD contributed 2 recordings each, and 9 children with ASD contributed 3 recordings each, while 26 and 17 TD children contributed respectively 2 and 3 recordings each) (see Figure S3 for illustration of the available recordings). This sample was employed to derive trajectories of visual exploration over the childhood years using mixed models analysis and considering both within-subject and between-subject effects (Mancini et al., 2019; Mutlu et al., 2013) and sliding windows approach (Sandini et al., 2018) (further detailed in the Method details subsection).

One-year follow-up longitudinal sample

In order to obtain purely longitudinal measure of change in visual exploration we used a smaller subsample that included children who had recordings obtained a year apart. Of note, 4 recordings out of 38 in the ASD group and 2 out of 26 available repeated recordings in the TD group were not included in this subsample as they were acquired after a period longer than 1 year (see Figure S3). Thus, this final longitudinal sample included 34 males with ASD (3.5 ± 1.2 years old) and 24 age-matched TD peers (3.3 ± 1.3 years old) who were evaluated a year later .

Stimuli and apparatus

The current experiment consisted of free-viewing of one episode of French cartoon "Trotro" lasting 2'53" (Lezoray, 2013). This cartoon was the first stimulus in an experiment that comprised simultaneous acquisition of High-density EEG recorded with a 129-channel Hydrocel Geodesic Sensor Net® (Electrical Geodesics Inc., Eugene, OR, USA). The findings concerning the EEG data are published separately (Jan et al., 2019). This cartoon depicts human-like interactions between three donkey characters at a relatively slow pace. The original soundtrack was preserved during recording. Gaze data were collected using Tobii TX300 eye tracker (https://www.tobiipro.com), sampled at 300Hz, with the exception of five recordings acquired at lower sampling frequency (60Hz) using Tobii TXL60. Screen size was identical for both machines (height: 1200 pixels (29°38') width: 1920 pixels (45°53')), with a refresh rate of 60Hz. Participants were seated at approximately 60 cm from the recording screen. The cartoon frames subtended visual angle of 26°47'×45°53' (height×width). A five-point calibration procedure consisting of child-friendly animations was performed using an inbuilt program in Tobii system. Upon verification, the calibration procedure was repeated if the eye-tracking

device failed to accurately detect participant's gaze position. The testing room had no windows and lighting conditions were held constant for all acquisitions.

Quantification and Statistical Analysis

Eye-tracking analysis

We excluded data from participants who showed poor screen attendance, defined as binocular gaze detection on less than 65% of frames of the video. To extract fixations we used the Tobii IV-T Fixation filter (Olsen, 2012) (i.e.,Velocity threshold: 30°/s; Velocity window length: 20ms. The adjacent fixations were merged (Maximum time between fixations was 75ms; Maximum angle between fixations was 0.5°).

Determining the "norm" of visual exploration

To define the normative gaze distribution ("norm"), against which we will compare the gaze data from the ASD group, we employed the kernel density distribution estimation function on gaze data from TD individuals on each frame of the video. Normative sample comprised 35 typically developing children (3.78 ± 1.3 years). To create normative gaze distribution we opted for a non-fixed bandwidth of a kernel as the gaze distribution characteristics vary importantly from frame to frame. Precisely, fixed bandwidth would result in over-smoothing the data in the mode and under-smoothing of extreme distribution cases of gaze data at tales. We used the state-of-the-art adaptive kernel density estimation that takes into account local characteristics of the data by employing adaptive kernel bandwidth (Botev, Grotowski, and Kroese, 2010). Thus a Gaussian kernel of an adaptive bandwidth was applied at each pair of gaze coordinates and the results were summed up to obtain an estimation of the density of gaze data (see Figure 1). Obtained density estimation reflects a probability of gaze allocation at the given location of the visual scene for a given group. This probability is higher at taller peaks (tightly packed kernels) of the distribution and diminishes toward the edges. We used the Matlab inbuilt function *contour* in order to delimit isolines of the gaze density matrix.

Quantifying the divergence in visual exploration

Upon the "norm" definition we calculated the distance of gaze data from this norm on each frame for each child with ASD ($n = 59$; 3.74 ± 1.4 years). Comparison to this normative pattern yielded a measure of *Proximity Index-PI* (see Figure 1). For the sake of comparison and interpretation, Proximity Index values were normalized (ranging from 0 to 1). A high Proximity Index value (closer to the mode of the density distribution) indicates that the visual exploration of the individual for a given frame is less divergent from the norm (more TD-like). A summary measure of divergence in visual exploration from the TD group was obtained by averaging the Proximity Index values for the total duration of the video.

Gaze patterns and association with behavioral data

To examine relation of this eye-tracking derived measure and overall functioning of children with ASD several correlational analyses were performed. PI was correlated with the measure of severity of symptoms of autism (ADOS) (Esler et al., 2015; Gotham et al., 2007) as well as with VABS-II Adaptive standard score (Sparrow, Balla, and Cicchetti, 2005) and Communication and Motor skills developmental quotients from PEP-3 (Schopler, 2005). In addition to these broader domains of adaptive and developmental functioning we conducted correlational analyses across their sub-domains. Thus we correlated Proximity Index with Communication (and the three subdomains composing it: Receptive, Expressive and Written language scale), Daily living skills, Socialization and Motor Domain of Vineland. Similarly, in the context of the developmental assessment we conducted correlation with scales composing Communication (Verbal & Preverbal Cognition, Receptive Language and Expressive Language scale) and Motor domain (Imitation, Fine Motor, Gross Motor scale).

Maturational changes in visual exploration of complex social scene

Beside understanding the behavioral correlates of atypical visual exploration in ASD we wanted to further characterize the developmental pathway of visual exploration of complex social scene in both groups of children. We opted for a sliding window approach adapted from Sandini et al. (2018) to delineate fine-grained changes in visual exploration on a group level. Available recordings from our unstructured longitudinal sample were first ordered according to the age in both groups separately. Then, a window encompassing 16 recordings was progressively moved across two groups, starting from the first 16 recordings in the youngest subjects until reaching the end part of the recording span for both groups. The choice of the window width was constrained by the sample size of our TD group. The longitudinal visits in our cohort are spaced a year from each other and the choice of a bigger window would result in important data loss in our group of TD children as the windows were skipped if they contained more than one recording from the same subject. The chosen window width yielded 40 sliding windows in both groups that were age-matched and spanned the period from 2.2 - 4.5 years old on average. Upon the creation of sliding windows, and to characterize the group visual behavior and its change with time, gaze data from TD group were pooled together to define the norm in each of the 40 age windows. In order to characterize the group visual behavior in ASD group, we performed the same by pooling the gaze data together to obtain

the normative distribution specific to this group in each window separately (see Figure S3). For this we applied kernel density distribution estimation function (Botev, Grotowski, and Kroese, 2010) on all 80 windows (40 in each group). Then we compared the relative gaze dispersion between groups on the estimated gaze density in each group and in each window separately. As a first measure of dispersion, we used the Median Contour Surface (MCS) plotted over density estimation function from the gaze data (see Figure 3 for illustration of this measure for a given frame). Greater values of this measure indicated more dispersed group gaze pattern (the group focus delimited by this median isoline on the given frame covered larger area). The additional measure of dispersion, Convergence Index (CI) was defined as the norm of the second derivative of the estimated density function for a given group $||f''||$, where $||.||$ is the Euclidean norm. Higher values of the Convergence Index corresponded to more focused, more coherent group gaze pattern (indicative of a more concave-down distribution). Both measures of dispersion were available on a frame-by-frame basis for each group.
To precisely characterize the change in visual exploration over a one year period values of the MCS as well as the CI were constrained to a smaller one-year follow-up longitudinal sample drawn from this larger sample used for sliding window analysis.

Proximity Index with regards to the visual properties of the animated scene

Considering the finding of enhanced sensitivity to the low level (pixel-level) saliency properties in ASD while watching static stimuli (Wang et al., 2015) our aim was to establish the level of influence that these basic visual properties exerted on gaze allocation in both groups while viewing animated social scene. As we used complex animated stimuli, basic visual properties are not expected to have a predominant role in directing gaze deployment, given the importance that the top-down attention influences exert in these (Itti, 2005; Itti and Koch, 2001; Itti et al., 2001). Nevertheless, our aim was to capture any discernible group difference in the influence of these bottom-up processes on gaze allocation during viewing of the animated stimuli. Moreover, we wanted to understand which of the low-level visual properties would contribute the most to spatio-temporal gaze allocation in both groups. To extract values of basic visual qualities of the scene, we used a salience model that has been extensively characterized in the literature (Itti and Koch, 2000; Itti, Koch, and Niebur, 1998; Itti et al., 2001; Koch and Ullman, 1985). This model extracts features based on simulated neurons in visual cortex: Color contrast (red/green and blue/yellow), Intensity contrast (light/dark), four orientation direction (0°, 45°, 90°, 135°), flicker (light offset and onset) and four motion energies (up, down, left, right) (Itti and Koch, 2001; Itti, Koch, and Niebur, 1998). The final saliency map is a result of the linear combination of these separate "channels" (Itti et al., 2001) into a unique scalar saliency map that guides attention. In order to disentangle the relative importance of the channels besides using the global conspicuity map we also used the channels taken separately (see Figure 5). We used two version of this model, an Itti and Koch model (I&K) (Koch and Ullman, 1985) and graph-based visual saliency (GBVS) model (Harel, Koch, and Perona, 2007).
The computations were performed at University of Geneva on the Baobab cluster.

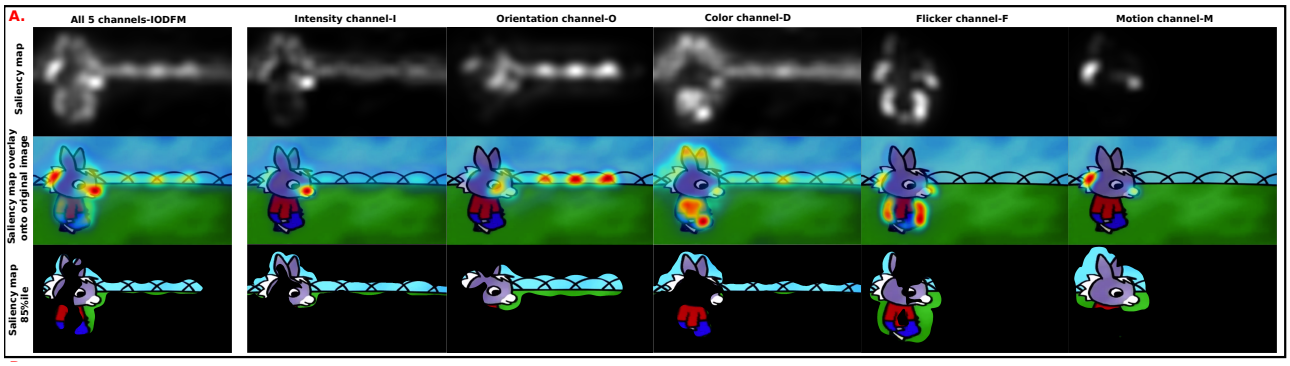

Figure 7: Illustration of the salience model. From left to right: full saliency model with all five channels combined and channels taken separately I-intensity, O-orientation, D-color, F-flicker and M-motion channel;
From top to bottom: A. Saliency map extracted for a given frame, Saliency map overlay on the original image, Original image with 15 % most salient parts shown

Intensity of social content and visual complexity

Furthermore, given the findings of the important impact of intensity of social content on social attention in ASD (Chita-Tegmark, 2016b; Frank, Vul, and Saxe, 2012), we aimed to test the hypothesis that the Proximity Index values will be lower for the moments in the videos with enhanced social complexity, involving two or three characters compared to moments involving only one character. However, with increasing number of characters the scene is inevitably richer in details. To control for the pure visual complexity aspects we calculated the length of edges delimiting image elements and used this measure as the proxy for visual complexity. Edge extraction was done on every image of the video using Canny method (Canny, 1986) implemented in Matlab (version 2017a; Mathworks, Natick, MA). This method finds edges by looking for local maxima of the gradient of intensity. The gradient is obtained using the derivative of a Gaussian

filter and uses two thresholds in order to detect strong and weak edges. Weak edges are retained only if connected to strong edges which makes this method relatively immune to noise.

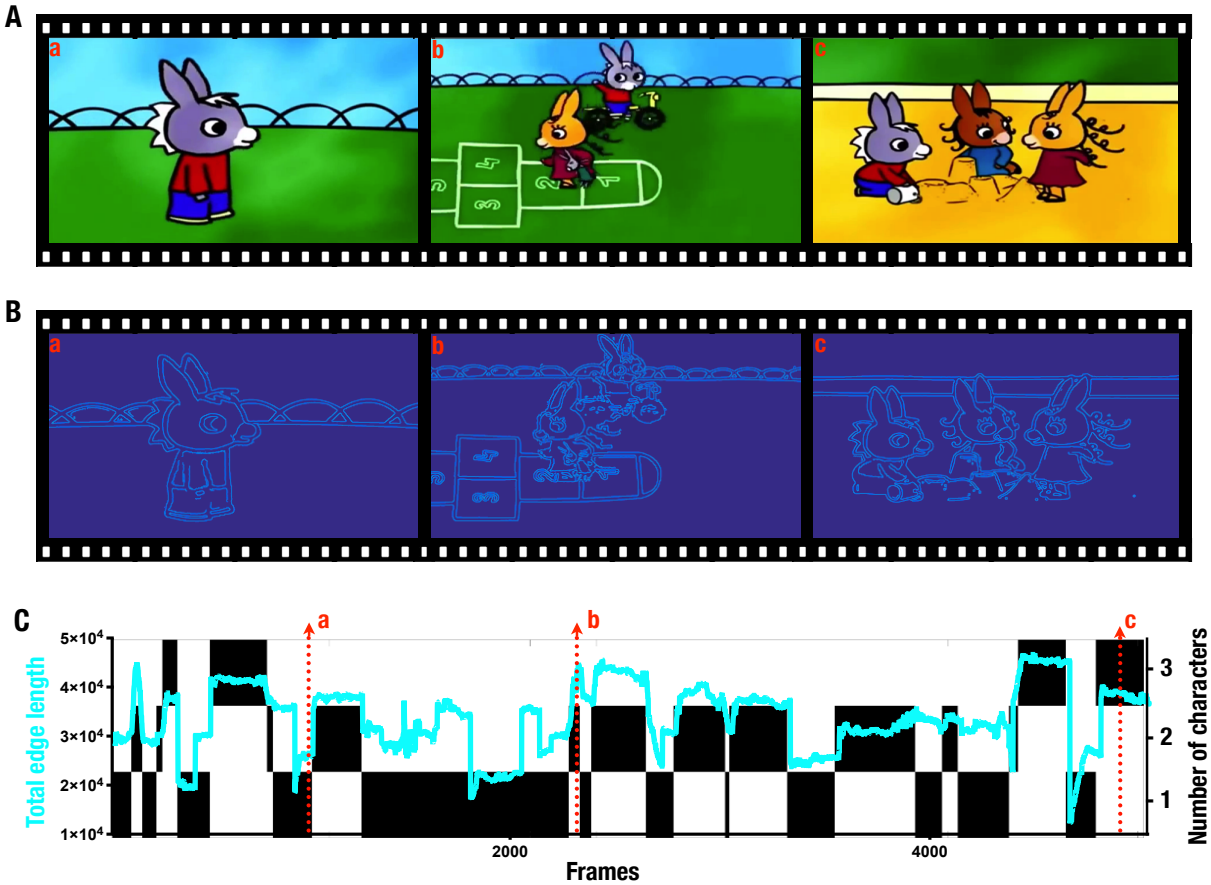

Figure 8: Illustration of the measures of social intensity and visual complexity. A. Three frames (denoted as a, b, c) illustrate three levels of social intensity; B. Visual complexity depicted using the edges of the images detected using the Canny method (Canny, 1986) for the frames a, b and c. C. Three social intensity levels are depicted in black for the duration of the video and corresponding visual complexity values (total edge length) are overlayed in cian.

Author Contributions

Conceptualization: N.K and M.S.; Methodology: N.K., C.S. and D.Z; Software: N.K. ;Formal Analysis: N.K. ; Investigation: N.K., M.F., H.F.S., R.K.J. ; Data Curation: N.K. ; Writing-Original Draft: N.K.; Writing-Review & Editing: All authors ; Visualization: N.K.; Supervision: M.S.; Project Administration: M.S.; Funding Acquisition: M.S.

Acknowledgments

We express our utmost gratitude to all families that took part in this study. We thank Stephanie Badoud, Isaline Humair, Léa Chambaz, Matthieu Mansion, Alexandra Bastos, Flore Couty, Myriam Speller, Sonia Richetin, François Robain, Lylia Ben Hadid, Marie-Agnès Graf, Sara Maglio, Irène Pittet for their precious contribution to data collection. Authors express their gratitude to Prof. Zdravko Botev for the precious help with Convergence Index calculation.

Funding for this study was provided by National Centre of Competence in Research (NCCR) Synapsy, financed by the Swiss National Science Foundation (SNF, Grant Number: 51AU40_125759), by SNF grants to M.S. (#163859 and #190084), and by the Fondation Pôle Autisme (https://www.pole-autisme.ch).

Declaration of Interests

The authors declare no competing interests.

References

American Psychiatric Association (May 22, 2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub. 1679 pp. ISBN: 978-0-89042-557-2.

Amso, Dima et al. (Mar. 2014). "Bottom-up attention orienting in young children with autism". In: *Journal of Autism and Developmental Disorders* 44.3, pp. 664–673. ISSN: 1573-3432. DOI: 10.1007/s10803-013-1925-5.

Avni, Inbar et al. (2019). "Children with autism observe social interactions in an idiosyncratic manner". In: *Autism Research* n/a (n/a). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aur.2234. ISSN: 1939-3806. DOI: 10.1002/aur.2234.

Bolton, Thomas A. W. et al. (Aug. 1, 2020). "Neural responses in autism during movie watching: Inter-individual response variability co-varies with symptomatology". In: *NeuroImage* 216, p. 116571. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2020.116571.

Bolton, Thomas A.W. et al. (Mar. 5, 2018). "Brain dynamics in ASD during movie-watching show idiosyncratic functional integration and segregation". In: *Human Brain Mapping* 39.6, pp. 2391–2404. ISSN: 1065-9471. DOI: 10.1002/hbm.24009.

Botev, Z. I., J. F. Grotowski, and D. P. Kroese (2010). "Kernel density estimation via diffusion". In: *Annals of Statistics* 38.5. ISBN: 0090-5364, pp. 2916–2957. ISSN: 00905364. DOI: 10.1214/10-AOS799.

Byrge, Lisa et al. (Apr. 8, 2015). "Idiosyncratic Brain Activation Patterns Are Associated with Poor Social Comprehension in Autism". In: *Journal of Neuroscience* 35.14. Publisher: Society for Neuroscience Section: Articles, pp. 5837–5850. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.5182-14.2015.

Canny, John (1986). "A computational approach to edge detection". In: *Ieee Transactions on Pattern Analysis and Machine Inligence*.

Chawarska, Katarzyna, Suzanne Macari, and Frederick Shic (2012). "Context modulates attention to social scenes in toddlers with autism". In: *Journal of Child Psychology and Psychiatry* 53.8, pp. 903–913. ISSN: 1469-7610. DOI: 10.1111/j.1469-7610.2012.02538.x.

Chawarska, Katarzyna and Frederick Shic (Dec. 2009). "Looking But Not Seeing: Atypical Visual Scanning and Recognition of Faces in 2 and 4-Year-Old Children with Autism Spectrum Disorder". In: *Journal of autism and developmental disorders* 39.12, pp. 1663–1672. ISSN: 0162-3257. DOI: 10.1007/s10803-009-0803-7.

Chevallier, Coralie et al. (Apr. 2012). "The social motivation theory of autism". In: *Trends in Cognitive Sciences* 16.4, pp. 231–239. ISSN: 1879-307X. DOI: 10.1016/j.tics.2012.02.007.

Chita-Tegmark, Meia (Sept. 1, 2016a). "Attention Allocation in ASD: a Review and Meta-analysis of Eye-Tracking Studies". In: *Review Journal of Autism and Developmental Disorders* 3.3, pp. 209–223. ISSN: 2195-7185. DOI: 10.1007/s40489-016-0077-x.

— (Jan. 1, 2016b). "Social attention in ASD: A review and meta-analysis of eye-tracking studies". In: *Research in Developmental Disabilities* 48, pp. 79–93. ISSN: 0891-4222. DOI: 10.1016/j.ridd.2015.10.011.

Colombo, John (2001). "The Development of Visual Attention in Infancy". In: *Annual Review of Psychology* 52.1. _eprint: https://doi.org/10.1146/annurev.psych.52.1.337, pp. 337–367. DOI: 10.1146/annurev.psych.52.1.337.

Constantino, John N. et al. (2017). "Infant viewing of social scenes is under genetic control and is atypical in autism". In: *Nature* 547.7663, pp. 340–344. ISSN: 1476-4687. DOI: 10.1038/nature22999.

Dawson, G. et al. (Dec. 1998). "Children with autism fail to orient to naturally occurring social stimuli". In: *Journal of Autism and Developmental Disorders* 28.6, pp. 479–485. ISSN: 0162-3257.

Dawson, Geraldine, Sara Jane Webb, and James McPartland (2005). "Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies". In: *Developmental Neuropsychology* 27.3, pp. 403–424. ISSN: 8756-5641. DOI: 10.1207/s15326942dn2703_6.

Dawson, Geraldine et al. (Jan. 2010). "Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model". In: *Pediatrics* 125.1, e17–23. ISSN: 1098-4275. DOI: 10.1542/peds.2009-0958.

Eisenberg, Michelle L. and Jeffrey M. Zacks (Jan. 1, 2016). "Ambient and focal visual processing of naturalistic activity". In: *Journal of Vision* 16.2. Publisher: The Association for Research in Vision and Ophthalmology, pp. 5–5. ISSN: 1534-7362. DOI: 10.1167/16.2.5.

Esler, Amy N. et al. (2015). "The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores". In: *Journal of Autism and Developmental Disorders* 45.9. Publisher: Springer US ISBN: 0162-3257, pp. 2704–2720. ISSN: 15733432. DOI: 10.1007/s10803-015-2432-7.

Falck-Ytter, Terje, Sven Bölte, and Gustaf Gredebäck (2013). "Eye tracking in early autism research." In: *Journal of neurodevelopmental disorders* 5.1. ISBN: 1866-1947 (Print)\r1866-1947 (Linking), p. 28. ISSN: 1866-1947. DOI: 10.1186/1866-1955-5-28.

Falck-Ytter, Terje and Claes von Hofsten (2011). "How special is social looking in ASD: a review". In: *Progress in Brain Research* 189, pp. 209–222. ISSN: 1875-7855. DOI: 10.1016/B978-0-444-53884-0.00026-9.

Falck-Ytter, Terje et al. (Oct. 2013). "Visualization and analysis of eye movement data from children with typical and atypical development". In: *Journal of Autism and Developmental Disorders* 43.10, pp. 2249–2258. ISSN: 1573-3432. DOI: 10.1007/s10803-013-1776-0.

Farber, D. A. and T. G. Beteleva (Sept. 1, 2005). "Formation of the System of Visual Perception in Ontogeny". In: *Human Physiology* 31.5, pp. 515–524. ISSN: 1608-3164. DOI: 10.1007/s10747-005-0091-3.

Franchak, John M. et al. (Apr. 1, 2016). "Free Viewing Gaze Behavior in Infants and Adults". In: *Infancy: The Official Journal of the International Society on Infant Studies* 21.3, pp. 262–287. ISSN: 1525-0008. DOI: 10.1111/infa.12119.

Franchini, Martina et al. (Mar. 1, 2017). "The effect of emotional intensity on responses to joint attention in preschoolers with an autism spectrum disorder". In: *Research in Autism Spectrum Disorders* 35, pp. 13–24. ISSN: 1750-9467. DOI: 10.1016/j.rasd.2016.11.010.

Franchini, Martina et al. (Aug. 1, 2018). "Early Adaptive Functioning Trajectories in Preschoolers With Autism Spectrum Disorders". In: *Journal of Pediatric Psychology* 43.7, pp. 800–813. ISSN: 0146-8693, 1465-735X. DOI: 10.1093/jpepsy/jsy024.

Frank, Michael C., Dima Amso, and Scott P. Johnson (Feb. 2014). "Visual search and attention to faces during early infancy". In: *Journal of Experimental Child Psychology* 118, pp. 13–26. ISSN: 1096-0457. DOI: 10.1016/j.jecp.2013.08.012.

Frank, Michael C., Edward Vul, and Scott P. Johnson (Feb. 2009). "Development of infants' attention to faces during the first year". In: *Cognition* 110.2, pp. 160–170. ISSN: 0010-0277. DOI: 10.1016/j.cognition.2008.11.010.

Frank, Michael C., Edward Vul, and Rebecca Saxe (2012). "Measuring the Development of Social Attention Using Free-Viewing". In: *Infancy* 17.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1532-7078.2011.00086.x, pp. 355–375. ISSN: 1532-7078. DOI: 10.1111/j.1532-7078.2011.00086.x.

Genovese, Christopher R., Nicole A. Lazar, and Thomas Nichols (Apr. 2002). "Thresholding of statistical maps in functional neuroimaging using the false discovery rate". In: *NeuroImage* 15.4, pp. 870–878. ISSN: 1053-8119. DOI: 10.1006/nimg.2001.1037.

Goren, C. C., M. Sarty, and P. Y. Wu (Oct. 1975). "Visual following and pattern discrimination of face-like stimuli by newborn infants". In: *Pediatrics* 56.4, pp. 544–549. ISSN: 0031-4005.

Gotham, Katherine et al. (2007). "The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity". In: *Journal of Autism and Developmental Disorders* 37.4. ISBN: 0162-3257, pp. 613–627. ISSN: 01623257. DOI: 10.1007/s10803-006-0280-1.

Green, David M. and John A. Swets (1966). *Signal detection theory and psychophysics*. Signal detection theory and psychophysics. Pages: xi, 455. Oxford, England: John Wiley. xi, 455.

Guillon, Quentin et al. (May 2014). "Visual social attention in autism spectrum disorder: insights from eye tracking studies". In: *Neuroscience and Biobehavioral Reviews* 42, pp. 279–297. ISSN: 1873-7528. DOI: 10.1016/j.neubiorev.2014.03.013.

Hanley, Mary et al. (Nov. 2013). "Spontaneous attention to faces in Asperger syndrome using ecologically valid static stimuli". In: *Autism: The International Journal of Research and Practice* 17.6, pp. 754–761. ISSN: 1461-7005. DOI: 10.1177/1362361312456746.

Harel, Jonathan, Christof Koch, and Pietro Perona (2007). "Graph-Based Visual Saliency". In: *Advances in Neural Information Processing Systems 19*. Ed. by B. Schölkopf, J. C. Platt, and T. Hoffman. MIT Press, pp. 545–552.

Hasson, Uri et al. (2009). "Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions". In: *Autism Research* 2.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aur.89 pp. 220–231. ISSN: 1939-3806. DOI: 10.1002/aur.89.

Helo, Andrea et al. (Oct. 2014). "The maturation of eye movement behavior: scene viewing characteristics in children and adults". In: *Vision Research* 103, pp. 83–91. ISSN: 1878-5646. DOI: 10.1016/j.visres.2014.08.006.

Helo, Andrea et al. (2016). "Eye movement patterns and visual attention during scene viewing in 3- to 12-month-olds". In: *Visual Neuroscience* 33, E014. ISSN: 1469-8714. DOI: 10.1017/S0952523816000110.

Helo, Andrea et al. (2017). "Influence of semantic consistency and perceptual features on visual attention during scene viewing in toddlers". In: *Infant Behavior & Development* 49, pp. 248–266. ISSN: 1934-8800. DOI: 10.1016/j.infbeh.2017.09.008.

Hendry, Alexandra et al. (Jan. 29, 2018). "Developmental change in look durations predicts later effortful control in toddlers at familial risk for ASD". In: *Journal of Neurodevelopmental Disorders* 10.1, p. 3. ISSN: 1866-1955. DOI: 10.1186/s11689-017-9219-4.

Hunnius, Sabine and Reint H. Geuze (2004). "Developmental Changes in Visual Scanning of dynamic faces and abstract stimuli in infants: A longitudinal study". In: *Infancy* 6.2. Place: US Publisher: Lawrence Erlbaum, pp. 231–255. ISSN: 1532-7078(Electronic),1525-0008(Print). DOI: 10.1207/s15327078in0602_5.

Hus Bal, Vanessa et al. (Oct. 2015). "Daily living skills in individuals with autism spectrum disorder from 2 to 21 years of age". In: *Autism : the international journal of research and practice* 19.7, pp. 774–784. ISSN: 1362-3613. DOI: 10.1177/1362361315575840.

Ito, Junji et al. (Apr. 24, 2017). "Switch from ambient to focal processing mode explains the dynamics of free viewing eye movements". In: *Scientific Reports* 7.1. Number: 1 Publisher: Nature Publishing Group, p. 1082. ISSN: 2045-2322. DOI: 10.1038/s41598-017-01076-w.

Itti, Laurent (2005). "Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes". In: *Visual Cognition* 12.6, pp. 1093–1123. ISSN: 1464-0716(Electronic),1350-6285(Print). DOI: 10.1080/13506280444000661.

Itti, Laurent and Christof Koch (2000). "A saliency-based search mechanism for overt and covert shifts of visual attention". In: *Vision Research* 40.10. ISBN: 0042-6989, pp. 1489–1506. ISSN: 00426989. DOI: 10.1016/S0042-6989(99)00163-7.

— (2001). "Feature combination strategies for saliency-based visual attention systems". In: *J. Electronic Imaging* 10, pp. 161–169. DOI: 10.1117/1.1333677.

Itti, Laurent, Christof Koch, and Ernst Niebur (1998). "A model of saliency-based visual attention for rapid scene analysis". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.11, pp. 1254–1259. ISSN: 0162-8828. DOI: 10.1109/34.730558.

Itti, Laurent et al. (2001). "Computational modelling of visual attention." In: *Nature reviews. Neuroscience* 2.3. ISBN: 1471-0048, pp. 194–203. ISSN: 1471-003X. DOI: 10.1038/35058500.

Jan, Reem K. et al. (2019). "Neural Processing of Dynamic Animated Social Interactions in Young Children With Autism Spectrum Disorder: A High-Density Electroencephalography Study". In: *Frontiers in Psychiatry* 10. ISSN: 1664-0640. DOI: 10.3389/fpsyt.2019.00582.

Johnson, Mark H. (July 2001). "Functional brain development in humans". In: *Nature Reviews Neuroscience* 2.7. Number: 7 Publisher: Nature Publishing Group, pp. 475–483. ISSN: 1471-0048. DOI: 10.1038/35081509.

Johnson, Mark H. et al. (Aug. 1, 1991). "Newborns' preferential tracking of face-like stimuli and its subsequent decline". In: *Cognition* 40.1, pp. 1–19. ISSN: 0010-0277. DOI: 10.1016/0010-0277(91)90045-6.

Jones, Warren and Ami Klin (Dec. 19, 2013). "Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism". In: *Nature* 504.7480, pp. 427–431. ISSN: 1476-4687. DOI: 10.1038/nature12715.

Kennedy, Daniel P. et al. (Nov. 20, 2017). "Genetic influence on eye movements to complex scenes at short timescales". In: *Current biology : CB* 27.22, 3554–3560.e3. ISSN: 0960-9822. DOI: 10.1016/j.cub.2017.10.007.

Kirkorian, Heather L., Daniel R. Anderson, and Rachel Keen (Apr. 2012). "Age differences in online processing of video: an eye movement study". In: *Child Development* 83.2, pp. 497–507. ISSN: 1467-8624. DOI: 10.1111/j.1467-8624.2011.01719.x.

Klin, Ami, Sarah Shultz, and Warren Jones (Mar. 2015). "Social visual engagement in infants and toddlers with autism: early developmental transitions and a model of pathogenesis". In: *Neuroscience and Biobehavioral Reviews* 50, pp. 189–203. ISSN: 1873-7528. DOI: 10.1016/j.neubiorev.2014.10.006.

Klin, Ami et al. (Sept. 1, 2002). "Visual Fixation Patterns During Viewing of Naturalistic Social Situations as Predictors of Social Competence in Individuals With Autism". In: *Archives of General Psychiatry* 59.9, pp. 809–816. ISSN: 0003-990X. DOI: 10.1001/archpsyc.59.9.809.

Klin, Ami et al. (Feb. 28, 2003). "The enactive mind, or from actions to cognition: lessons from autism." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 358.1430, pp. 345–360. ISSN: 0962-8436. DOI: 10.1098/rstb.2002.1202.

Klin, Ami et al. (Apr. 1, 2007). "Social and Communication Abilities and Disabilities in Higher Functioning Individuals with Autism Spectrum Disorders: The Vineland and the ADOS". In: *Journal of Autism and Developmental Disorders* 37.4, pp. 748–759. ISSN: 1573-3432. DOI: 10.1007/s10803-006-0229-4.

Klin, Ami et al. (May 14, 2009). "Two-year-olds with autism orient to nonsocial contingencies rather than biological motion". In: *Nature* 459.7244, pp. 257–261. ISSN: 0028-0836. DOI: 10.1038/nature07868.

Koch, C. and S. Ullman (1985). "Shifts in selective visual attention: towards the underlying neural circuitry". In: *Human Neurobiology* 4.4, pp. 219–227. ISSN: 0721-9075.

Kojovic, Nada et al. (Oct. 2019). "Sensory Processing Issues and Their Association with Social Difficulties in Children with Autism Spectrum Disorders". In: *Journal of Clinical Medicine* 8.10, p. 1508. DOI: 10.3390/jcm8101508.

Lezoray, S. (2013). *Trotro est amoureux [Cartoon]*. In collab. with E. Cases.

Lord, C. et al. (June 2000). "The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism". In: *Journal of Autism and Developmental Disorders* 30.3, pp. 205–223. ISSN: 0162-3257.

Lord, Catherine et al. (2012). *Autism diagnostic observation schedule: ADOS-2*. OCLC: 851410387. Los Angeles, Calif.: Western Psychological Services.

Mancini, Valentina et al. (June 4, 2019). "Positive psychotic symptoms are associated with divergent developmental trajectories of hippocampal volume during late adolescence in patients with 22q11DS". In: *Molecular Psychiatry*. ISSN: 1359-4184, 1476-5578. DOI: 10.1038/s41380-019-0443-z.

Mutlu, A. Kadir et al. (Nov. 15, 2013). "Sex differences in thickness, and folding developments throughout the cortex". In: *NeuroImage* 82, pp. 200–207. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2013.05.076.

Nakano, Tamami et al. (Oct. 7, 2010). "Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour". In: *Proceedings of the Royal Society B: Biological Sciences* 277.1696, pp. 2935–2943. ISSN: 0962-8452. DOI: 10.1098/rspb.2010.0587.

Olsen, Anneli (2012). "The Tobii I-VT Fixation Filter". In: p. 21.

Parish-Morris, Julia et al. (Apr. 13, 2019). "Adaptation to different communicative contexts: an eye tracking study of autistic adults". In: *Journal of Neurodevelopmental Disorders* 11.1, p. 5. ISSN: 1866-1955. DOI: 10.1186/s11689-019-9265-1.

Riby, D. and P. J. B. Hancock (Feb. 2009). "Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism". In: *Journal of intellectual disability research: JIDR* 53.2, pp. 169–181. ISSN: 1365-2788. DOI: 10.1111/j.1365-2788.2008.01142.x.

Richler, Jennifer et al. (2010). "Developmental Trajectories of Restricted and Repetitive Behaviors and Interests in Children with Autism Spectrum Disorders". In: *Development and psychopathology* 22.1, pp. 55–69. ISSN: 0954-5794. DOI: 10.1017/S0954579409990265.

Rider, Andrew T. et al. (Feb. 1, 2018). "Semantic content outweighs low-level saliency in determining children's and adults' fixation of movies". In: *Journal of Experimental Child Psychology* 166, pp. 293–309. ISSN: 0022-0965. DOI: 10.1016/j.jecp.2017.09.002.

Rosa Salva, Orsola et al. (Apr. 20, 2011). "The Evolution of Social Orienting: Evidence from Chicks (Gallus gallus) and Human Newborns". In: *PLoS ONE* 6.4. Ed. by Gonzalo García de Polavieja, e18802. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0018802.

Rosenthal, Robert (1991). *Meta-Analytic Procedures for Social Science Research*. Vol. 15. Publication Title: Educational Researcher Issue: 8 ISSN: 0013189X. 18 pp. ISBN: 0-8039-4245-1 0-8039-4246-X. DOI: 10.2307/1175262.

Sandini, Corrado et al. (May 18, 2018). "Development of Structural Covariance From Childhood to Adolescence: A Longitudinal Study in 22q11.2DS". In: *Frontiers in Neuroscience* 12. ISSN: 1662-4548. DOI: 10.3389/fnins.2018.00327.

Sasson, Noah J. et al. (2008). "Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays". In: *Autism Research* 1.1, pp. 31–42. ISSN: 1939-3806. DOI: 10.1002/aur.4.

Sasson, Noah J. et al. (Feb. 2011). "Brief report: Circumscribed attention in young children with autism". In: *Journal of Autism and Developmental Disorders* 41.2, pp. 242–247. ISSN: 1573-3432. DOI: 10.1007/s10803-010-1038-3.

Schopler, Eric (2005). *PEP-3: Psychoeducational Profile*. Google-Books-ID: L1C1XwAACAAJ. PRO-ED. Dimensions 30.0x25.0 cm, book.

Schultz, Robert T. (May 2005). "Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area". In: *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience* 23.2, pp. 125–141. ISSN: 0736-5748. DOI: 10.1016/j.ijdevneu.2004.12.012.

Shic, Frederick et al. (July 2007). "Measuring context: The gaze patterns of children with autism evaluated from the bottom-up". In: *2007 IEEE 6th International Conference on Development and Learning*. 2007 IEEE 6th International Conference on Development and Learning. ISSN: 2161-9476, pp. 70–75. DOI: 10.1109/DEVLRN.2007.4354067.

Shic, Frederick et al. (Aug. 2008). "Autism, eye-tracking, entropy". In: *2008 7th IEEE International Conference on Development and Learning*. 2008 7th IEEE International Conference on Development and Learning. ISSN: 2161-9476, pp. 73–78. DOI: 10.1109/DEVLRN.2008.4640808.

Simion, Francesca, Lucia Regolin, and Hermann Bulf (Jan. 15, 2008). "A predisposition for biological motion in the newborn baby". In: *Proceedings of the National Academy of Sciences* 105.2. Publisher: National Academy of Sciences Section: Biological Sciences, pp. 809–813. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0707021105.

Simion, Francesca et al. (2001). "The origins of face perception: specific versus non-specific mechanisms". In: *Infant and Child Development* 10.1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/icd.247, pp. 59–65. ISSN: 1522-7219. DOI: 10.1002/icd.247.

Sparrow, Sara S., David Balla, and Domenic V. Cicchetti (2005). *Vineland II: Vineland Adaptative Behavior Scales : Survey Forms Manual : a Revision of Hte Vineland Social Maturity Scale by Edgar A. Doll*. Pearson. 330 pp.

Speer, Leslie L. et al. (May 2007). "Face processing in children with autism: effects of stimulus contents and type". In: *Autism: The International Journal of Research and Practice* 11.3, pp. 265–277. ISSN: 1362-3613. DOI: 10.1177/1362361307076925.

Sperdin, Holger Franz et al. (Feb. 27, 2018). "Early alterations of social brain networks in young children with autism". In: *eLife* 7. Ed. by Jan-Marino Ramirez, e31670. ISSN: 2050-084X. DOI: 10.7554/eLife.31670.

Unema, Pieter J. A. et al. (Apr. 1, 2005). "Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration". In: *Visual Cognition* 12.3. Publisher: Routledge _eprint: https://doi.org/10.1080/13506280444000409, pp. 473–494. ISSN: 1350-6285. DOI: 10.1080/13506280444000409.

Valenza, E. et al. (Aug. 1996). "Face preference at birth". In: *Journal of Experimental Psychology. Human Perception and Performance* 22.4, pp. 892–903. ISSN: 0096-1523. DOI: 10.1037//0096-1523.22.4.892.

Wang, Quan et al. (2018). "Operationalizing atypical gaze in toddlers with autism spectrum disorders: a cohesion-based approach". In: *Molecular Autism* 9, p. 25. ISSN: 2040-2392. DOI: 10.1186/s13229-018-0211-y.

Wang, Shuo et al. (2015). "Atypical Visual Saliency in Autism Spectrum Disorder Quantified through Model-Based Eye Tracking". In: *Neuron* 88.3. Publisher: Elsevier Ltd ISBN: 9780124201187, pp. 604–616. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.09.042.

Yarbus, Alfred L. (1967). *Eye movements and vision*. New York: Plenum.

Young, Gregory S. et al. (Sept. 2009). "Gaze behavior and affect at 6 months: predicting clinical outcomes and language development in typically developing infants and infants at risk for autism". In: *Developmental Science* 12.5, pp. 798–814. ISSN: 1467-7687. DOI: 10.1111/j.1467-7687.2009.00833.x.