

Supplemental Material S1. Operation definitions and calculation method for the dependent variables.

Dependent variable	Operational definition	Calculation method
a. Mean Total Time Spent Fixating on the Photograph across All Elements	It reflects overall level of visual attention to a photograph.	It sums the durations of all the fixations on anywhere within the visual stimuli during the 5 s photograph display time.
b. Mean Percentage of Each Participant's Own Time Spent Fixating on the Human	It reflects how human figures within a photograph may attract and maintain visual attention, in comparison with other visual elements, such as trees, animals, clocks, etc.	It is calculated by dividing participant's viewing time on human figures by his or her own viewing time on the screen.
c. Ratio of Time Spent Fixating on the Human Figures Relative to Its Size	It reflects strength of the visual attraction to human figures. Ratios over 1 mean that the participants spent more time fixated on the human figures than would be expected based on the human figures' size (see below an explanation of the ratio measure).	It is calculated by dividing the actual time spent on the human figures by the expected amount of time spent the human figures based on their size alone.
d. Ratio of Time Spent Fixating on the Head Relative to its Size	It reflects strength of the visual attraction to head. Ratios over 1 mean that the participants spent more time fixated on the head than would be expected based on the head's size.	It is calculated by dividing the actual time spent on the head by the expected amount of time spent the head based on its size alone.
e. Mean Absolute Latency of Fixation on the Human Figures	It reflects prioritized visual attention to the human figures relative to other visual elements within the whole photo.	It is the time elapsed between the onset of the presentation of the image and the first fixation to the human figures.
f. Mean Relative Latency of Fixation on the Human Figures	It reflects prioritized visual attention to the human element with the consideration of participant's own viewing pace.	It is the amount of time elapsed between each participant's first fixation anywhere on the image, and the first fixation on the human figures.
g. Percentage of Photographs in which Participants Fixated on the Human Figures	It reflects how likely that a photograph is to attract a participant's attention at all when a photograph contain human figures.	It is calculated by dividing the number of photographs in which each individual fixated at least once on the human figures by the total of 32 possible photographs.

Dependent Variable b. Mean Percentage of Each Participant's Own Time Spent Fixating on the Human

The examination of the total time spent fixating on the entire scene across is critical for analyzing the percentage of time fixated on the human figures. There are two ways to calculate the percentages of fixations on an area of interest. One method is to use the total possible viewing time (the stimulus display time) as the denominator. If screen display time is 5 s, group A spends 2 s viewing a human figure, and group B spends 3 s, then it is possible to find a between-group difference (group A vs. group B; 40% vs. 60%). A different approach is to set the denominator as the total time spent by the

participant him/herself. Thus, for instance, if group A spends only 3 s fixating on the whole screen and group B spends 5 s, then the between-group difference in the proportion of one's own viewing time spent viewing the human figure will be smaller (group A vs. group B; 60% vs. 67%), even though the groups differ in their absolute viewing time. Therefore, gaze behaviors can be deemed typical or atypical depending on the selection of denominator. A clear statement of percentage calculation is critical because the selection of denominator may affect between-group comparisons. The current study used participants' own viewing time as denominator for the percentage measure.

Dependent Variable c. Ratio of Time Spent Fixating on the Head Relative to its Size

Viewers do not usually fixate on every part of a visual scene and tend to fixate on informative parts of the scene (Rayner, 2009). The ratio measure is critical as we expect that visual fixations would be unequally divided among elements within the photograph. Some elements would be expected to receive fixation from only some participants or for only brief periods of time, whereas some other elements would be expected to receive more fixations than expected based on their size alone. For example, if an element occupies 10% of the space but is viewed for 30% of the viewing time, then this element is receiving more attention than would be expected based on its size; if an element occupies 30% of the space, but is viewed 10% of the viewing time, then this element receives less attention than would be expected based on its size (Fletcher-Watson, Findlay, Leekam, & Benson, 2008; Wilkinson & Light, 2011). Therefore, it was valuable not only to evaluate how likely an element attracted attention (measure (a) and the percentage of photographs in which human figures were fixated), but also how well the element maintained attention (measures (d) and (e)) within the 5 s viewing period.

Dependent Variable d. Ratio of Time Spent Fixating on the Head Relative to its Size

In addition to fixations to the eyes, other body parts may also be critical for social understanding and communication (Yu & Smith, 2016). For individuals with typical development (TD), if social saliency is stronger in the heads or eyes than other parts of the body, one might expect that visual preference to the whole body differs from that to the heads or eyes. Given the social profile, the relation might not hold in individuals with autism spectrum disorder (ASD). Therefore, in examining the strength of the visual attraction, we not only evaluated the ratio of time spent in fixating on the whole body (measure (c)), but also on the head alone (measure (d)).

Dependent Variable g: Percentage of Photographs in which Participants Fixated on the Human Figures

This measure was calculated by dividing the number of photographs in which each individual fixated at least once on the human figures by the total of 32 possible photographs. The below table indicates that the TD group fixated on the human figures in most photographs (99%), as did the Down syndrome (DS) group (97%). In contrast, the ASD group had a lower rate (83%). Omnibus independent-samples Kruskal–Wallis tests revealed significant group differences, $\chi^2(2) = 23.063, p < .001$. Pair-wise independent-samples Mann–Whitney *U* tests indicated that the ASD group viewed lower percentages of photographs than the DS group, $U = 9.00, p = .001$, and also lower percentages of

photographs than the TD group, $U = 6.00, p < .001$. The DS group viewed significantly lower percentages of photographs than the TD group, $U = 18.00, p = .015$.

Table S1. Percentage of photographs in which participants fixated on the human figures.

Etiology	Participant code	Total number of photos with fixations on the human figures	Percentage
ASD			
	ASD1	29	91%
	ASD2	32	100%
	ASD3	25	78%
	ASD4	28	88%
	ASD5	25	78%
	ASD6	27	84%
	ASD7	27	84%
	ASD8	23	72%
	ASD9	29	91%
	ASD10	22	69%
	Mean	26.7	83%
DS			
	DS1	32	100%
	DS2	31	97%
	DS3	31	97%
	DS4	31	97%
	DS5	31	97%
	DS6	30	94%
	DS7	31	97%
	DS8	30	94%
	DS9	16*	100%
	DS10	31	97%
	Mean	29.4	97%
TD			
	TD1	32	100%
	TD2	32	100%
	TD3	31	97%
	TD4	32	100%
	TD5	32	100%
	TD6	31	97%
	TD7	32	100%
	TD8	32	100%
	TD9	32	100%
	TD10	32	100%
	Mean	31.8	99%

Note. TD = typical development; DS = Down syndrome; ASD = autism spectrum disorder.

*Participant DS9 only watched 16 photographs because the experimenter failed to provide the remaining 16 photographs during the session; therefore, percentages were calculated with a denominator of 16 for this participant.

References

- Fletcher-Watson, S., Leekam, S. R., Findlay, J. M., & Stanton, E. C. (2008). Brief report: Young adults with autism spectrum disorder show normal attention to eye-gaze information-evidence from a new change blindness paradigm. *Journal of Autism and Developmental Disorders*, *38*(9), 1785–1790.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.
- Wilkinson, K. M., & Light, J. (2011). Preliminary investigation of visual attention to human figures in photographs: Potential considerations for the design of aided AAC visual scene displays. *Journal of Speech, Language, and Hearing Research*, *54*(6), 1644–1657.
- Yu, C., & Smith, L. B. (2016). Multiple sensory-motor pathways lead to coordinated visual attention. *Cognitive Science*, *41*(S1), 5–31.