**Infants' looking preferences for social versus non-social objects reflect genetic variation
and are linked to later language development**

Ana Maria Portugal^{1*}, Charlotte Viktorsson², Mark J. Taylor³, Luke Mason⁴, Kristiina Tammimies^{1,5}, Angelica Ronald⁴, Terje Falck-Ytter^{1,2,6}

¹ Center of Neurodevelopmental Disorders (KIND), Centre for Psychiatry Research, Department of Women's and Children's Health, Karolinska Institutet & Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden

² Development and Neurodiversity Lab, Department of Psychology, Uppsala University, Uppsala, Sweden

³ Department of Medical Epidemiology & Biostatistics, Karolinska Institutet, Stockholm, Sweden

⁴ Department of Psychological Sciences, Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK

⁵ Astrid Lindgren Children's Hospital, Karolinska University Hospital, Region Stockholm, Sweden

⁶ Swedish Collegium for Advanced Study, Uppsala, Sweden

*Corresponding author. Email: ana.maria.portugal@ki.se

Running title: Face looking in infancy: heritability, specificity, and link to language and socio-communication

Word count: 2297 excluding abstract, methods, references, tables, and figures.

8 figures and tables

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

ABSTRACT 146/ 150

To what extent do individual differences in infants' early preference for faces versus non-facial objects reflect genetic and environmental factors? In a sample of 536 5-month-old same-sex twins, we assessed attention to faces using eye-tracking in two ways: initial orienting to faces at the start of the trial (thought to reflect subcortical processing) and sustained face preference throughout the trial (thought to reflect emerging attention control). We show that both face orienting and preference were heritable ($h^2 = .19$, CI: .04-.33; $h^2 = .46$, CI: .33-.57, respectively); while shared environment was statistically nonsignificant. Face preference was associated positively with later parent-reported verbal competence ($\beta = 0.14$, $p = .014$). This is the first study to show that individual differences in young infants' selection of perceptual input – social versus non-social – are heritable, providing a new developmental perspective on gene-environment interplay occurring at the level of eye movements.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Infants learn a lot from looking at other people, and from the subsequent interaction with them. At the same time, the developing infant is confronted with the massive task of learning about non-social objects and events. Optimally prioritizing between attending to social stimuli, such as faces, and the non-social world is not straightforward and is likely to reflect the infants' degree of social understanding and motivation as well as maturation of the cognitive system. Atypical attention to social versus non-social objects has been implicated in autism, a heritable neurodevelopmental condition partly defined by social communication difficulties ¹⁻⁴. It is increasingly acknowledged that social looking preferences represent a complex set of underlying processes ⁵, and there is large variation in the typical population with regards to social versus non-social visual preferences ^{3,6,7}. In line with this, recent data suggest that specific aspects of social preferences such as attention to eyes versus mouth of other people's faces are highly heritable in infants ⁸ and young children ⁹. Yet, no study has asked the more fundamental question of to what extent visual preference for faces versus non-social information in infancy reflects genetic variation in the population (which, at the extreme, could be associated with heritable clinical conditions such as autism ^{3,10,11}) or differences in environmental influences.

Faces are special; they selectively attract attention already at birth and they act as a learning platform for cognitive and social development ¹²⁻¹⁵. Early life biases to face orienting (i.e., fast first looks to face-like configurations in the periphery already present at birth) are proposed to be sub-cortically mediated whereas sustained preference to attend to faces requires later maturing cortical top-down structures ¹⁴. By 6 months of age, top-down control enables flexible looking behaviour permitting the baby to preferentially attend to the face but also to shift attention away from it and to the other stimuli in the environment ^{14,16,17}.

Here, we studied the genetic and environmental influences underlying individual differences in three early emerging aspects of selective attention to social versus non-social

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

information: orienting (first look to faces), sustained preference (ratio of looking time in the face), and visual information exploration (number of objects, social and non-social, looked at). For this, we used Babytwins Study Sweden (BATSS¹⁸), a Swedish community sample of dizygotic and monozygotic 5-months-olds twins, and gaze-based experimental measurements of looking to faces when presented together with other non-face objects. Given the links reported between visual attention to faces and autism¹⁻⁴ and between attention control and Attention-Deficit/Hyperactivity Disorders (ADHD¹⁰), we also tested whether the different emerging aspects of attention to faces early in life were associated with polygenic scores (common genetic variance) related to autism and ADHD, and with later autism and ADHD related outcomes (socio-communication, and self-regulation). Because greater attention to faces in infancy is also thought to predict better language outcomes later in life^{e.g.,19,20} we also studied whether the different aspects of attention to faces are associated with later language skills.

Some recent studies have studied genetic influences in gaze allocation in complex social and non-social scenes^{8,9,21}. Compared to these, our study was unique in that it employed a genetically informed and longitudinal approach to experimentally dissociate looking behaviour to social (faces) versus non-social objects at a very early period when bidirectional processes between the child and their environment (e.g., caregivers) have just started to emerge and might unfold into cascading effects in social and cognitive development.

RESULTS

Sample descriptive statistics are presented in **Table 1** and **2**. As expected, we found a face bias significantly above chance level, for both face orienting (proportion of trials with first look to the face; one-sample $t_{\text{Twin 1}}(273) = 8.98, p < .001$; $t_{\text{Twin 2}}(261) = 7.85, p < .001$)

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

and preference (proportion of time spent looking at the face; $t_{\text{Twin 1}}(273) = 28.29, p < .001$;

$t_{\text{Twin 2}}(261) = 26.31, p < .001$) – **Fig. 1**.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Table 1. Descriptive statistics of the primary face looking measures. Statistics presented as Mean (SD) / min-max. MZ, monozygotic; DZ, dizygotic.

	Overall	MZ Females	MZ Males	DZ Females	DZ Males	Skewness
N (twins)	536	135	158	116	127	
Age (in days)	168 (9) 145-203	168 (9) 153-194	167 (8) 150-187	167 (8) 153-189	168 (10) 145-203	0.62
# Valid trials	5.77 (0.52) 4-7	5.76 (0.52) 4-7	5.75 (0.55) 4-6	5.75 (0.56) 4-6	5.82 (0.44) 4-6	-2.12
% Missing gaze samples	.28 (.13) 0-.61	.28 (.13) .01-.61	.28 (.12) .05-.58	.29 (.13) .02-.60	.27 (.14) 0-.58	0.13
Face orienting (% first look to the face) *	.30 (.19) 0-1	.29 (.19) 0-.80	.27 (.2) 0-.83	0.34 (.19) 0-.83	.30 (.19) 0-1	0.51
Face preference (% on face) *	.44 (.14) .09-.81	.45 (.14) .11-.78	.42 (.15) .09-.80	.45 (.13) .11-.70	.43 (.14) .15-.81	0.14
# Objects explored (in 0-10 seconds, out of maximum 5)	3.64 (0.54) 2-4.83	3.61 (0.54) 2.17-4.83	3.66 (0.52) 2.17-4.83	3.60 (0.56) 2-4.83	3.69 (0.55) 2.25-4.67	-0.34

* Chance level of face looking was .20 (1 in 5 objects was a face)

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Table 2. Descriptive statistics of parent-report development measures at 14 and 24 months. Statistics presented as Mean (SD) / min-max.

	N (n girls)	Age	Overall	Females	Males	Skewness
ECBQ Self-regulation						
at 14 months	436 (196)	442 (20) 386-525	4.32 (0.58) 2.40-6.15	4.30 (0.59) 2.4-6.15	4.33 (0.56) 2.4-5.76	0.10
at 24 months	358 (180)	755 (26) 707-920	4.56 (0.65) 2.73-6.25	4.69 (0.61) 2.73-6	4.44 (0.66) 2.8-6.25	-0.17
ITC Social communication						
at 14 months	418 (196)	444 (20) 387-525	34.93 (6.85) 11-51	35.84 (6.09) 20-51	34.13 (7.38) 11-51	-0.4
CDI vocabulary						
at 14 months	420 (197)	443 (20) 386-502	82.52 (64.47) 1-332	92.90 (61.85) 1-305	73.35 (65.48) 3-332	1.16
at 24 months	335 (168)	756 (23) 707-920	207.62 (156.91) 1-689	251.66 (143.71) 18-627	163.32 (157.55) 1-689	0.72

ECBQ: Early Childhood Behavior Questionnaire

ITC: Infant Toddler Checklist

CDI: Communicative Development Inventory

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Figure 1. The face pop-out paradigm. A. Illustration of an infant viewing one trial of the face pop-out task. B. Distribution and boxplot of the three primary looking measures derived from the task, across 536 5-month-old infants: face orienting (chance level highlighted as a dashed vertical line), face preference (chance level highlighted as a dashed vertical line), and number of objects explored.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Higher twin correlations in MZ than DZ pairs suggested genetic influences on face orienting and face preference, and low correlations on face orienting suggested unique environment influences (**Table 3**). For the number of objects explored, there was very low similarity within both MZ and DZ, suggesting negligible familial effects. For this reason, further twin and genetic results were not reported for this measure (but see results in **Tables S1-2**).

Table 3. Twin correlations (95% confidence intervals are shown in parentheses) for the primary face looking measures, separate for MZ and DZ pairs.

	N (Twin pairs*)	Face orienting (% first look to the face)	Preference for face (% on face)	# Objects explored (in 0-10 seconds)
MZ	155	.20 [.03-.34]	.46 [.32-.57]	.05 [-.12-.22]
DZ	130	.05 [-.13-.23]	.21 [.01-.38]	.15 [-.03-.31]

** Incomplete twin pairs*

Correlations were derived from the twin models with age and sex included as covariates.

We proceeded to fit a bivariate twin model with face orienting and face preference (and age and sex included as covariates) to our data (but see all univariate results in **Tables S1-6**). All twin assumptions were met (**Table S7**). The phenotypic correlation between the face measures was positive and moderate ($r_{Ph} = .30$, CI: .22-.37). The CTCT correlations suggested genetic influences on the phenotypic correlation because the MZ CTCT correlation ($r_{CTCT\ MZ} = .21$, CI: .10-.31) was higher than the DZ CTCT correlation ($r_{CTCT\ DZ} = -.03$, CI: -.17-.11). Non-additive genetic influences were suggested because the DZ correlations were less than half the MZ correlations, so genetic variance needs to be interpreted as broad-sense heritability. The AE model (the model without shared environment influences) did not have a significantly poorer fit compared to the ACE model, so this was reported in **Table 4** and **Fig. 2** (but see ACE estimates in **Table S8**).

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

For face orienting there were small significant genetic effects ($A = .19$, CI: .04-.33). Face preference had a significant and higher heritability ($A = .46$, CI: .33-.57). The bivariate results showed that in terms of unique environment, virtually no E on face preference was shared with face orienting (98% of total E influencing face preference was unique to that variable). For genetic factors, the model estimates suggested that .16 (CI: .03-.51, of the total A of .46) of the genetic influences on face preference were shared with, and .29 (CI: 0-.45, of .46) were unique from, first look to faces – **Fig. 2**. A follow-up analysis testing two nested models constraining the shared or the unique influences showed that the unique genetic variance on face preference was not statistically significant (although marginal), whereas the shared genetic influences were – **Table 4**.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Table 4. Bivariate twin model fit statistics for face orienting and face preference. Best-fitting model in bold.

Model	# Parameters	-2LL	df	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value
Fully Sat	32	1151.37	1040	-928.63	NA	NA	NA	NA
ACE								
ACE	15	1167.82	1057	-946.18	Fully Sat.	16.45	17	.492
ACE nested models								
AE	12	1169.33	1060	-950.67	ACE	1.51	3	.680
CE	12	1175.34	1060	-944.66	ACE	7.52	3	.057
E	9	1211.82	1063	-914.18	ACE	44.00	6	< .001
AE nested models								
Unique path=0	11	1172.28	1061	-949.72	AE	2.95	1	.086
Shared path=0	11	1179.41	1061	-942.59	AE	10.08	1	.001

Model definitions. The Fully Sat. model is the fully saturated model of the observed data, which models the means and variances for both variables, and the phenotypic and cross-twin-cross-trait correlations between the two variables, separately for each twin in a pair and across zygosity. The best-fitting model (in bold) was the non-significant and most parsimonious model, as well as the one with the lowest AIC.

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Figure 2. *Schematic AE bivariate twin model for face orienting and face preference. A = genetic influences, E = unique environment influences. The dotted line represents the shared genetic influences (significant), whereas the dashed line represents the unique genetic influences (marginally significant).*

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Associations with polygenic scores

The associations between the face looking measures (face orienting and face preference) and polygenic scores for autism, ADHD, schizophrenia, depression, and bipolar disorder were tested but none were statistically significant (see **Table S9**).

Longitudinal phenotypic associations

We tested the associations between the looking measures and subsequent parent-report measures of language, socio-communication, and self-regulation – **Table 5**. We found a statistically significant positive association between preference for the face and receptive vocabulary (comprehension in the CDI) at 14 months.

Further, a non-significant (p -value below .05 but not below the pre-registered FDR threshold) positive association was found between the number of objects explored and the child's expressive vocabulary (production in the CDI) at 24 months, as well as non-significant positive associations between face preference and number of objects explored and socio-communication skills at 14 months.

Table 5. Results of multiple GEE analyses with 14 months and 24 months parent-report measures as outcome variables, age, sex, and the looking measures measured at 5 months as predictors, and twin pair id as cluster-defining variable. For each model, all predictors were entered together; hence statistics represent unique contributions for each predictor.

Significant predictors are in bold (see p -threshold in the False Discovery Rate threshold column).

ECBQ Self-regulation at 14 months	Beta	Uncorrected p-value	FDR Threshold
Age (in days, scaled)	-0.08	.159	0.01
Sex (reference level: Female)	0.06	.608	0.04
Face orienting (% first look to the face)	0.07	.215	0.02
Preference for face (% on face)	0.00	.935	0.05
# Objects explored (in 0-10 seconds)	0.04	.377	0.03

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

ECBQ Self-regulation at 24 months	Beta	Uncorrected <i>p</i>-value	FDR Threshold
Age (in days, scaled)	0.05	0.475	0.03
Sex (reference level: Female)	-0.38	0.003	0.01
Face orienting (% first look to the face)	0.06	0.248	0.02
Face preference (% on face)	0.04	0.548	0.05
# Objects explored (in 0-10 seconds)	0.03	0.530	0.04
ITC Social communication at 14 months	Beta	Uncorrected <i>p</i>-value	FDR Threshold
Age (in days, scaled)	0.07	.348	.04
Sex (reference level: Female)	-0.27	.030	.02
Face orienting (% first look to the face)	-0.01	.918	.05
Face preference (% on face)	0.13	.026	.01
# Objects explored (in 0-10 seconds)	0.10	.031	.03
CDI Receptive vocabulary at 14 months	Beta	Uncorrected <i>p</i>-value	FDR Threshold
Age (in days, scaled)	0.15	.007	.01
Sex (reference level: Female)	-0.34	.008	.02
Face orienting (% first look to the face)	-0.01	.809	.05
Face preference (% on face)	0.14	.014	.03
# Objects explored (in 0-10 seconds)	0.07	.142	.04
CDI Expressive vocabulary at 24 months	Beta	Uncorrected <i>p</i>-value	FDR Threshold
Age (in days, scaled)	0.21	< .001	.02
Sex (reference level: Female)	-0.58	< .001	.01
Face orienting (% first look to the face)	-0.01	.872	.05
Face preference (% on face)	0.04	.550	.04
# Objects explored (in 0-10 seconds)	0.10	.046	.03

DISCUSSION

Attention preferences to social information, and looking to faces more specifically, has been suggested to play a key role in the development of social interaction and cognition²². Against this background, it is striking that we found substantial variability in infants' attention to faces at an early point in life – before brain systems supporting social communication are fully developed. Further, this variability was heritable for both face orienting at the beginning of the task and sustained face preference. In contrast, shared environment appeared to play a negligible role for infants' tendency to preferentially attend to faces versus non-social objects. The pattern of these results supports the view that, already during infancy, there is genetic variability to curate one's visual environment^{8,9,21}, at least in terms of looking preferences for one or the other. This can be seen as a type of gene-environment correlation²³ appearing very early in life: an active exposure to different visual environments, even at the subtle level of face versus non-social objects, provides access to

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

very different learning opportunities (active gene-environment correlation) as well as potentially evokes different reactions from the environment (evocative gene-environment correlation, e.g., from the caregiver). Visual environment selection by means of selective attention (under genetic influence) is one of the first ways infants can actively create or constrain their own visual experience and social interactions, emerging before other exploratory behaviours such as pointing, grasping, or crawling to targets or partners. It is very likely this phenomenon has significant influences over time, and indeed in our study we showed an association between face looking and later developmental outcomes.

Face preference, indexed by the looking time to the face relative to looking time to all objects in the scene, had a heritability of 46%. For face orienting, indexed by the proportion of first looks to the face, we showed heritability of 19%. This held even when we considered other related measures (i.e., latency to the face, see **Appendix S2**). Given the adaptive and survival value of face orienting, it is not surprising that there is limited genetic variation linked to this phenotype. It is plausible to think that the genetic variation we found might be related to individual social motivation and interests (endogenous, cortical attention control), to general attention abilities, or both. This should be investigated in future studies.

For face and non-face objects exploration, indexed by the number of objects an infant looked at, we did not find significant evidence for familial effects (genetic or shared environment), and variability was best explained solely by unique environmental factors (which include measurement error). Perhaps our sample was too young to be displaying a stable measure of exploration, as indeed exploratory gaze patterns have been shown to be less consistent in infancy ²⁴, and/or our study was underpowered to detect subtle familial influences in this case.

First looks to faces and face preference were moderately correlated, and there was evidence for some shared genetic influences on face preference from face orienting. While

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

initial orienting and sustained attention to faces are hypothesized to be dissociated in terms of underlying brain networks (subcortical vs cortical networks¹⁴), it is possible that this shared variance might be driven by shared sub-cortical processes influencing both phenotypes, by early emerging face-specific cortical structures influencing both face preference and orienting²⁵, or by an inflation of the co-variance due to potential dependency of the measures (where you look first will probably influence to some extent your preferences at longer timescales). These hypotheses are not mutually exclusive, and they should be disentangled in future studies. At the same time, the bivariate model estimates indicated that most (two-thirds) of the genetic influences on preference for face looking were unique to this phenotype, which could indicate that sustained attention to the face had indeed a dissociated genetic etiology from orienting to the face. However, the unique genetic factors on face preference were statistically non-significant (although marginally significant). This can be a power issue, but we cannot completely rule out that all genetic factors contributing to variance in preference for faces are shared with face orienting. There are likely both shared and unique genetic influences on face preference, but we need a bigger sample to demonstrate these effects.

Face orienting was not associated with later developmental outcomes, which is in line with previous research that face orienting is intact across the lifespan and in typical and atypical development²⁶. Face preference was associated with higher infant's receptive vocabulary at 14 months, which is related to findings from the same project that showed a positive relationship between eye-versus-mouth looking and vocabulary⁸. In addition, a non-statistically (i.e., p -value below .05 but above the pre-registered FDR threshold) positive association was found between the number of objects explored and the child's expressive vocabulary at 24 months; and between face preference and number of objects explored and socio-communication skills at 14 months. Although these findings need to be replicated, they could indicate that face preference and attention control in infancy is associated with long-

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

term socio-communication and language, which are in line with reports of atypical engagement with social scenes in infants that later develop autism ¹¹.

We did not find any statistically significant associations between the looking measures and self-regulation, although self-regulation early in life is thought to be supported by co-regulation from the caregiver and by developing attention control ²⁷. Future studies should aim to use more fine-grained measures of self-regulation, and potentially more stable measures of social attention control, to investigate the long-term links between these two. Further, we did not find any associations between looking measures and genome-wide polygenic scores for autism or ADHD. While this might be reflecting true null effects, it is possible that the current polygenic scores do not have yet enough predictive power to detect these links.

The study has some notable limitations. First, the number of face pop-out trials in our study was 6 – increasing the number of trials could potentially lead to more stable measures of infant face orienting and objects exploration (though likely at a cost of increased participant attrition). Second, while the BATSS study included almost 30% of the same-sex twin population, it reflected families with a higher socioeconomic status compared to Stockholm normative ¹⁸. This needs to be considered in the generalizability of our results as genetic and environmental estimates may vary in samples where SES has a wider distribution e.g.,²⁸.

In conclusion, our findings inform us about the etiological influences on attention preferences for social input early in life, and its developmental associations in the first 2 years of life. These results suggest two forms of gene-environment interplay unfolding at a micro-level in infancy. Firstly, because selective attention influences the input received, heritable preferences in infancy can be seen as a biological-driven selection of the environment.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Further, looking at faces may cause reactions from the social partner. Both will have cascading effects in cognitive and social development.

METHODS

Three hundred and eleven families of same-sex twins were recruited to the BabyTwins Study Sweden (BATSS ¹⁸) and invited for an initial in-person assessment at 5 months at Karolinska Institutet (data collection from April 2016 to February 2020), and multiple follow-up online questionnaires at 14 months, 24 months, and 36 months (ongoing data collection). Parents gave informed consent to take part at each time point. BATSS was approved by the Regional Ethical Review Board in Stockholm and was conducted in accordance with the Declaration of Helsinki. The main project sample description and inclusion criteria are described elsewhere ¹⁸. Zygosity was estimated based on DNA sampled from all infants.

For the current study, a total of 28 twins were excluded due to: parent-reported twin-to-twin transfusion syndrome, report of seizures at the time of birth, very low birth weight (<1.5 Kg), or report of spina bifida. In addition, 23 infants did not complete the eye-tracking assessment due to technical reasons, time constraints, bad calibration, or tiredness. Of the infants that completed this session, 35 infants did not provide enough valid trials, so they were excluded as well (see below the criteria for valid trials). The final sample consisted of 536 individuals (285 pairs with at least one twin with data, of those 251 pairs had complete data). There were no statistically significant differences between the excluded (based on invalid trials or not completion of eye-tracking assessment) and included infants in terms of either parental education level, family income, sex, or age.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Eye-tracking Protocol

To record infants' gaze a Tobii TX300 eye-tracker was used (sampling rate of 120 Hz) with Matlab and Psychtoolbox (for stimuli presentation) with custom algorithms. An initial 5-point calibration was called before the start of the task battery. This eye-tracking battery involved rotations of free-viewing of the face pop-out task (see below) and dynamic scenes trials (mixture of social and abstract content), gaze-contingent gap-overlap trials, pupillary light reflex measurements, and post-calibration sequences; and lasted for about 10 minutes.

The face pop-out task

The face pop-out task^{3,4,29} was used to measure the various attentional processes involved in visual attention. It consisted of the presentation (fixed order, 20 seconds each) of a set of 6 different complex displays of objects including a face (with direct eye-gaze; 3 male, 3 female; counterbalancing ethnicity and location of the face within the array) and 4 non-face competitors (including a 'noise' stimulus generated from the same face, a mobile phone, a bird, and a car) – see **Fig. 3**. This type of display and the relatively long presentation time allow the study of the variation in active seeking of social information and attention flexibility by estimating the timing and preferential attention to faces and non-faces stimuli.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Figure 3. Scheme of the face pop-out experimental stimuli.

Before each slide, a small animation was presented and gaze-contingent methods started the presentation of the displays synchronous to the infant's look to the central animation, ensuring the gaze was at the centre of the screen at the start of the trial. One infant viewed 7 trials of the face pop-out (the first trial was repeated because the protocol had to be restarted) – all trials were included in the analysis.

Computation of primary measures

Gaze data for each pop-out trial was processed after data collection using custom-written Matlab scripts (see analysis steps and Areas of Interest, AOIs, in **Appendix S1** and **Fig. S1**). Any trials with a proportion of valid data (non-missing) less than .25 (25%), with a total duration of data for a trial less than 5 seconds, or where no look to an AOI was made, were excluded. Measures were averaged across trials for each infant if at least 4 valid trials were found. The distributions and boxplots for the proportion of valid trials where each AOI was the first AOI looked at and the proportion of looking time to each AOI (relative to all AOIs) can be seen in **Appendix S1 Fig. S2**.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Face initial orienting was operationalized as the proportion of first looks to the face (first look to faces, i.e., the number of valid trials where the face was the first object looked at in relation to the number of valid trials), in deviation of the pre-registered plan of using a composite measure of the proportion of first look to faces and the mean latency to look at the face. This decision was based on a modest correlation found between latency and proportion of first look to the face (Standardized $\beta = -.33$, $p < .001$) and unmet twin assumptions for the composite measure (driven by unmet equality of means and variances across zygosity and twin order for latency). However, sensitivity analyses with this composite led to a similar pattern of genetic univariate and bivariate findings (see **Appendix S2 Tables S10-S13**).

Face preference was operationalized as the mean ratio of looking at the face, i.e., the sum of looking time at the face AOI divided by the sum of looking time at all AOIs averaged across valid trials.

Object exploration was operationalized as the mean number of objects looked at, averaged across valid trials. Each array of objects was presented for 20 seconds, a longer duration than in previous studies' protocols^{3,4,30} where shorter versions were used (12 and 15 seconds). The longer duration meant that it was likely that infants looked at the 5 objects during the trial (in 45% of all trials, the 5 objects were looked at). For this reason, and in deviation of the pre-registered plan, object exploration was estimated based on only the first 10 seconds of the trial (in 22% of all trials, the 5 objects were looked at). The cut-off did not influence the results (see **Appendix S3 Tables S14-S16**).

Gaze quality measures

To control for potential effects of gaze quality in analyses, we estimated two gaze quality variables: the average proportion of missing data in the task (operationalized as the

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

ratio of missing gaze per total data collected, averaged across valid trials) and the number of valid trials.

Genome-wide polygenic scores

DNA samples were genotyped using Infinium Global Screening Array (Illumina, San Diego, CA, USA). Genotype quality control and processing were done using standard procedures and are described elsewhere¹⁸. Polygenic scores were calculated based on the most recent and largest genome-wide association studies for ADHD³¹, autism³², bipolar disorder³³, major depressive disorder³⁴, and schizophrenia³⁵. The polygenic scores were calculated using the PRS-CS (polygenic prediction via Bayesian regression and continuous shrinkage priors) method³⁶. For the polygenic score analysis, the first 10 principal components of ancestry were included as covariates.

Parent-rated developmental questionnaires

Social communication

The CSBS DP Infant Toddler Checklist (ITC³⁷) was used to measure socio-communicative behaviours (as indexed by the total score) at 14 months. A lower score is indicative of communication difficulties.

Language

The Swedish Early Communicative Development Inventory (CDI^{38,39}) adapted from the Macarthur-Bates Communicative Development Inventory), was used to measure vocabulary at 14 (the Words and Gestures form) and 24 months (the Words and Sentences form). Given prior findings on another study with the same sample⁸, we used receptive vocabulary (number of words the child understands) at 14 months; and we used expressive vocabulary (number of words the child understands and says) at 24 months.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

Self-regulation

The Early Childhood Behavior Questionnaire (ECBQ ⁴⁰) was used to measure self-regulation (as indexed by the effortful control scale of the questionnaires) at 14 (short-form, 107 items) and 24 (very-short-form, 36 items) months.

Analysis plans

An analysis plan for this study was registered in OSF (osf.io/5fbvk) prior to data cleaning and analysis. R software (version 4.0.0) was used for all data computation and analyses.

For twin models, the OpenMx package ⁴¹ (version 2.18.1 with NPSOL optimizer) with full-information maximum likelihood estimation was used, which allows for partially complete pairs (one twin missing) to be included. Twin models work by comparing the level of within-pair similarity separately for monozygotic twins (MZ; who share 100% of their segregating genetic material) and dizygotic twins (DZ; who on average share 50%).

Univariate twin models estimate the relative contribution of genetic and environmental factors to the variation in a phenotype, by comparing the correlation between twins; while bivariate twin models further estimate the relative contributions of genetic and environmental factors to the covariation between two phenotypes, by comparing cross-trait cross-twin (CTCTs) correlations, i.e., the correlation between one phenotype for one twin and another phenotype for their co-twin. The variation or covariation can be decomposed into additive genetic influences (A, heritability, which increases twin similarity), non-shared environment (E, environmental influences that differ between twins and decrease twin similarity, including measurement error), and either shared environment (C, environmental influences that increase twin similarity regardless of zygosity, e.g., family socioeconomic status) or non-

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

additive genetic (D) effects (D and C variance cannot be estimated simultaneously from twin data alone).

For each looking measure, both saturated models (which test for the assumptions of equality of mean and variances across twin order and zygosity) and univariate twin models were fitted separately and reported in **Tables S1-S6**. A bivariate saturated model (including test for the assumption of the equality of phenotypic and CTCTs correlations across twin order and zygosity) and a bivariate twin model were fitted between phenotypes that were phenotypically correlated and shown to be heritable from twin correlations (i.e., first look to faces and face preference). A Cholesky decomposition was used to examine genetic influences on face preference that were either unique to face preference or shared with first look to faces. The best fitting Cholesky decomposition was reported based on the AIC fit statistic (Akaike information criterion, the lowest value corresponds to the best model), and on being parsimonious (fewer number of parameters estimated) and non-significance (meaning that there was no decrement in fit compared to the saturated or the genetic model, indexed by the χ^2 distribution). Twin and CTCTs correlations were derived from the constrained saturated models, in which means, variances, phenotypic and CTCTs correlations were constrained to be equal across twin order and zygosity. When the pattern of correlations suggested non-additive genetic effects (MZ correlation more than twice the DZ correlation) a decision was made to report an ACE model rather than an ADE model to our data due to sample size (but see Table S15 for the ADE bivariate model results).

Association analyses were conducted, whenever possible, with the whole twin sample (i.e., including both twins in a pair, including pairs with one twin missing) using linear regression models implemented as generalized estimating equations (GEE, using the *drgee*

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

package ⁴²), with cluster-robust SE to account for family relatedness, to derive Beta estimates and p-values. All measures were scaled so that Beta estimates are standardized.

Age (in days) and sex were always included in twin models and added to the GEE models as covariates. Associations between the gaze quality covariates (proportion missing gaze and number of valid trials) and the gaze-based primary visual attention measures were tested within the GEE framework (one linear model with both covariates as predictors were run for each primary variable). If statistically significant, the gaze quality measures were regressed out from the main dependent variables before all other analyses.

Acknowledgements

The authors thank all participating families, as well as research assistants Linnea Hamrefors, Joy Hättstrand, Lynnea, Myers, Johanna Kronqvist, Sofia Jönsson, Anna Kernell, Carolin Schreiner, Sophie Lingö, Angelinn Liljebäck, Isabelle Enedahl, Matthis Andreasson, Lisa Belfrage, Mattias Savallampi, Isabelle Ocklind, Hjalmar Nobel Norrman, and Ingrid Schragge. Special thank you to Danyang Li for contributing to the genetic data processing.

This research was supported by Stiftelsen Riksbankens Jubileumsfond (NHS14-1802:1; Pro Futura Scientia [in collaboration with SCAS]); the Swedish Research Council (2018-06232); Knut and Alice Wallenberg foundation; the European Union (EU-MSCA Initial Training Networks: 642996, BRAINVIEW; 814302, SAPIENS). K.T. was supported by the Swedish Foundation for Strategic Research FFL18-0104. L.M. has received funding from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement No. 115300, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007—2013) and EFPIA companies' in kind

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

contribution; the Innovative Medicines Initiative 2 Joint Undertaking under Grant Agreement No. 777394. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI; awards from the Medical Research Council (MR/K021389/1; MR/T003057/1). The authors acknowledge the KI Biobank for handling the biological samples, SNP&SEQ Technology Platform, Uppsala University, for genotyping, and the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973" for computations. Any views expressed are those of the author(s) and not necessarily those of the funders.

Conception and design: A.M.P., T.F.Y., A.R., and K.T.; data preparation: A.M.P., L.M., and K.T.; analysis of data: A.M.P. with contributions of T.F.Y., M.T.; writing (original draft): A.M.P., T.F.Y., M.T., C.V.; writing (critical review and editing): all authors.

The authors have declared that they have no competing or potential conflicts of interest.

Data and code availability section

The Face Pop-out preprocessing workflow (implemented in Matlab and described in Appendix S1) is part of a shared agreement and available from L.M. on a reasonable request. The statistical analysis scripts (implemented in R) are publicly available in OSF (https://osf.io/zseh2/?view_only=1d7a815ff87148a6af5a6b58c427419c). Relevant data are available from the corresponding author on a reasonable request. Note that sharing of pseudonymized (coded) data will require a data processor agreement (DPA) according to Swedish and EU law.

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

REFERENCES

1. Nyström, P., Bölte, S. & Falck-Ytter, T. Responding to Other People's Direct Gaze: Alterations in Gaze Behavior in Infants at Risk for Autism Occur on Very Short Timescales. *J Autism Dev Disord* 47, 3498–3509 (2017).
2. Gangi, D. N. *et al.* Declining Gaze to Faces in Infants Developing Autism Spectrum Disorder: Evidence From Two Independent Cohorts. *Child Dev* 92, e285–e295 (2021).
3. Hendry, A. *et al.* Developmental change in look durations predicts later effortful control in toddlers at familial risk for ASD. *J Neurodev Disord* 10, 3 (2018).
4. Elsabbagh, M. *et al.* The development of face orienting mechanisms in infants at-risk for autism. *Behav Brain Res* 251, 147–154 (2013).
5. Salley, B. & Colombo, J. Conceptualizing Social Attention in Developmental Research. *Soc Dev* 25, 687–703 (2016).
6. Pierce, K. *et al.* Eye Tracking Reveals Abnormal Visual Preference for Geometric Images as an Early Biomarker of an Autism Spectrum Disorder Subtype Associated With Increased Symptom Severity. *Biol Psychiat* 79, 657–666 (2016).
7. Frank, M. C., Amso, D. & Johnson, S. P. Visual search and attention to faces during early infancy. *J Exp Child Psychol* 118, 13–26 (2013).
8. Viktorsson, C. *et al.* Infants' attention to eyes is an independent, heritable trait predicting later verbal competence. *Biorxiv* 2021.10.23.465424 (2021) doi:10.1101/2021.10.23.465424.
9. Constantino, J. N. *et al.* Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature* 547, 340–344 (2017).
10. Gui, A. *et al.* Look duration at the face as a developmental endophenotype: elucidating pathways to autism and ADHD. *Development and Psychopathology* (2020) doi:10.1017/s0954579420000930.
11. Chawarska, K., Macari, S. & Shic, F. Decreased Spontaneous Attention to Social Scenes in 6-Month-Old Infants Later Diagnosed with Autism Spectrum Disorders. *Biol Psychiat* 74, 195–203 (2013).
12. Amso, D. & Scerif, G. The attentive brain: insights from developmental cognitive neuroscience. *Nature reviews neuroscience* 16, 606--619 (2015).
13. Reynolds, G. D. & Roth, K. C. The Development of Attentional Biases for Faces in Infancy: A Developmental Systems Perspective. *Front Psychol* 09, 222 (2018).

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

14. Johnson, M. H., Senju, A. & Tomalski, P. The two-process theory of face processing: Modifications based on two decades of data from infants and adults. *Neurosci Biobehav Rev* 50, 169–179 (2015).
15. Simion, F. & Giorgio, E. D. Face perception and processing in early infancy: inborn predispositions and developmental changes. *Front Psychol* 6, 969 (2015).
16. Johnson, M. H., Dziurawiec, S., Ellis, H. & Morton, J. Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40, 1–19 (1991).
17. Gliga, T. & Csibra, G. Seeing the face through the eyes: a developmental perspective on face expertise. in *Progress in Brain Research* vol. 164 323–339 (2007).
18. Falck-Ytter, T. *et al.* Babytwins Study Sweden (BATSS): A multi-method infant twin study of genetic and environmental factors influencing infant brain and behavioral development. *Twin Research and Human Genetics* (2021).
19. Brooks, R. & Meltzoff, A. N. The development of gaze following and its relation to language. *Developmental Sci* 8, 535–543 (2005).
20. Morales, M., Mundy, P. & Rojas, J. Following the direction of gaze and language development in 6-month-olds. *Infant Behav Dev* 21, 373–377 (1998).
21. Kennedy, D. P. *et al.* Genetic Influence on Eye Movements to Complex Scenes at Short Timescales. *Curr Biol* 27, 3554–3560.e3 (2017).
22. Johnson, M. H. Subcortical face processing. *Nat Rev Neurosci* 6, 766–774 (2005).
23. Plomin, R., DeFries, J. C. & Loehlin, J. C. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychol Bull* 84, 309–322 (1977).
24. Franchak, J. M., Heeger, D. J., Hasson, U. & Adolph, K. E. Free Viewing Gaze Behavior in Infants and Adults. *Infancy* 21, 262–287 (2016).
25. Buiatti, M. *et al.* Cortical route for facelike pattern processing in human newborns. *Proc National Acad Sci* 116, 201812419 (2019).
26. Johnson, M. H. Autism: Demise of the Innate Social Orienting Hypothesis. *Curr Biol* 24, R30–R31 (2014).
27. Hendry, A., Jones, E. J. H. & Charman, T. Executive function in the first three years of life: Precursors, predictors and patterns. *Developmental Review* 42, 1–33 (2016).
28. Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B. & Gottesman, I. I. Socioeconomic Status Modifies Heritability of IQ in Young Children. *Psychol Sci* 14, 623–628 (2003).
29. Jones, E. J. H. *et al.* Eurosibs: Towards robust measurement of infant neurocognitive predictors of autism across Europe. *Infant Behav Dev* 57, 101316 (2019).

Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development by Portugal et al.

30. Gliga, T., Elsabbagh, M., Andravizou, A. & Johnson, M. Faces attract infants' attention in complex displays. *Infancy* 5, 550–562 (2009).

31. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* 51, 63–75 (2019).

32. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* 51, 431–444 (2019).

33. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 51, 793–803 (2019).

34. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci* 22, 343–352 (2019).

35. Consortium, T. S. W. G. of the P. G., Ripke, S., Walters, J. T. & O'Donovan, M. C. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *Medrxiv* 2020.09.12.20192922 (2020) doi:10.1101/2020.09.12.20192922.

36. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 10, 1776 (2019).

37. Wetherby, A. M., Allen, L., Cleary, J., Kublin, K. & Goldstein, H. Validity and Reliability of the Communication and Symbolic Behavior Scales Developmental Profile With Very Young Children. *J Speech Lang Hear Res* 45, 1202–1218 (2002).

38. Berglund, E. & Eriksson, M. Communicative Development in Swedish Children 16-28 months old: The Swedish Early Communicative Development Inventory—Words and Sentences. *Scand J Psychol* 41, 133–144 (2000).

39. Eriksson, M. & Berglund, E. Swedish early communicative development inventories: words and gestures. *First Lang* 19, 55–90 (1999).

40. Putnam, S. P., Gartstein, M. A. & Rothbart, M. K. Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behav Dev* 29, 386–401 (2006).

41. Neale, M. C. *et al.* OpenMx 2.0: Extended Structural Equation and Statistical Modeling. *Psychometrika* 81, 535–549 (2016).

42. Zetterqvist, J. & Sjölander, A. Doubly Robust Estimation with the R Package drgee. *Epidemiologic Methods* 4, 69–86 (2015).

Supplementary Information

for the research paper “Infants’ looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development” by Ana Maria Portugal et al.

Supplemental Tables

Table S 1. Univariate saturated model for the number of objects explored (0-10 seconds) including covariates.

Model	# Parameters	-2LL	df	AIC	Comparative fit with saturated model		
					$\Delta \chi^2$	Δdf	P value
Saturated	12	1507.21	524	459.21	-	-	-
Age	11	1507.21	525	457.21	0.01	1	.941
Sex	11	1509.15	525	459.15	1.94	1	.163
1.	10	1512.52	526	460.52	5.32	2	.070
2.	9	1512.53	527	458.53	5.32	3	.150
3.	7	1514.50	529	456.50	7.30	5	.200
4.	6	1514.93	530	454.93	7.72	6	.259

Model definitions. The baseline model is the fully saturated model of the observed data, which models the means and variances separately for each twin in a pair and across zygosity. **Age**. Testing the significance of age, **Sex**. Testing the significance of sex, **1.** Equating means across twins within a pair, **2.** Equating means across zygosity, **3.** Equating variances across twins within a pair, and **4.** Equating variances across zygosity (i.e., the constrained saturated model).

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Table S 2. Univariate twin model fit statistics and parameter estimates for number of objects explored (0-10 seconds). Best-fitting model in bold.

Model	# parameters	-2LL	df	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A	C	E
Fully Sat	12	1507.21	524	459.21	NA	-	-	-	-	-	-
ACE	6	1515.50	530	455.50	Fully Sat.	8.29	6	.218	0	.10	.90
AE	5	1516.41	531	454.41	ACE	0.91	1	.341	.10	0	.90
CE	5	1515.50	531	453.50	ACE	0	1	> .999	0	.10 (0-.22)	.90 (.78-1)
E	4	1518.00	532	454.00	ACE	2.50	2	.286	0	0	1

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

Table S 3. Univariate saturated model for face orienting (proportion of first look to faces) including covariates.

Model	# Parameters	-2LL	df	AIC	Comparative fit with saturated model		
					$\Delta \chi^2$	Δdf	P value
Saturated	12	-260.80	524	-1308.80	-	-	-
Age	11	-258.49	525	-1308.49	2.31	1	.129
Sex	11	-258.94	525	-1308.94	1.86	1	.173
1.	10	-260.14	526	-1312.14	0.66	2	.718
2.	9	-254.83	527	-1308.83	5.97	3	.113
3.	7	-252.42	529	-1310.42	8.38	5	.136
4.	6	-252.27	530	-1312.27	8.53	6	.202

Model definitions. The baseline model is the fully saturated model of the observed data, which models the means and variances separately for each twin in a pair and across zygosity. **Age**. Testing the significance of age, **Sex**. Testing the significance of sex, **1**. Equating means across twins within a pair, **2**. Equating means across zygosity, **3**. Equating variances across twins within a pair, and **4**. Equating variances across zygosity (i.e., the constrained saturated model).

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in $-2LL$ statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

Table S 4. Univariate twin model fit statistics and parameter estimates for face orienting (proportion of first look to faces). Best-fitting model in bold.

Model	# parameters	-2LL	df	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A	C/D	E
Fully Sat	12	-260.80	524	-1308.80	NA	-	-	-	-	-	-
ACE	6	-252.05	530	-1312.05	Fully Sat.	8.75	6	.188	.18	0	.82
AE	5	-252.05	531	-1314.05	ACE	0.00	1	> .999	.18 (.03-.32)	0	.82 (.68-.97)
CE	5	-250.87	531	-1312.87	ACE	1.18	1	.278	0	.13	.87
E	4	-246.42	532	-1310.42	ACE	5.63	2	.060	0	0	1
ADE	6	-252.27	530	-1312.27	Fully Sat.	8.53	6	.202	.01	.19	.81
AE	5	-252.05	531	-1314.05	ADE	0.22	1	.638	.18	0	.82
E	4	-246.42	532	-1310.42	ADE	5.85	2	.054	0	0	1

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in $-2LL$ statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

Table S 5. Univariate saturated model for the face preference (% looking time on face) including covariates.

Model	# Parameters	-2LL	df	AIC	Comparative fit with saturated model		
					$\Delta \chi^2$	Δdf	P value
Saturated	12	1463.71	524	415.71	-	-	-
Age	11	1470.19	525	420.19	6.48	1	.011
Sex	11	1466.57	525	416.57	2.86	1	.091

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

1.	10	1466.16	526	414.16	2.45	2	.294
2.	9	1466.17	527	412.17	2.46	3	.483
3.	7	1467.85	529	409.85	4.14	5	.529
4.	6	1468.96	530	408.96	5.25	6	.512

Model definitions. The baseline model is the fully saturated model of the observed data, which models the means and variances separately for each twin in a pair and across zygosity. **Age**. Testing the significance of age, **Sex**. Testing the significance of sex, **1**. Equating means across twins within a pair, **2**. Equating means across zygosity, **3**. Equating variances across twins within a pair, and **4**. Equating variances across zygosity (i.e., the constrained saturated model).

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

Table S 6. Univariate twin model fit statistics and parameter estimates for face preference (% looking time on face). Best-fitting model in bold.

Model	# parameters	-2LL	<i>df</i>	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A	C/D	E
Fully Sat	12	1463.71	524	415.71	NA	-	-	-	-	-	-
ACE	6	1469.00	530	409.00	Fully Sat.	5.29	6	.507	.45	0	.55
AE	5	1469.00	531	407.00	ACE	0.00	1	> .999	.45 (.33-.56)	0	.55 (.44-.68)
CE	5	1474.29	531	412.29	ACE	5.28	1	.022	0	.36	.64
E	4	1507.67	532	443.67	ACE	38.66	2	< .001	0	0	1
ADE	6	1468.96	530	408.96	Fully Sat.	5.25	6	.512	.37	.08	.54
AE	5	1469.00	531	407.00	ADE	0.05	1	.831	.45	0	.55
E	4	1507.67	532	443.67	ADE	38.71	2	< .001	0	0	1

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Table S 7. Assumption testing for the bivariate model between face orienting (proportion of first look to faces) and face preference.

Model	# Parameters	-2LL	df	AIC	Comparative fit with saturated model		
					$\Delta \chi^2$	Δdf	P value
Saturated	32	1151.37	1040	-928.63	-	-	-
5.	15	1165.79	1057	-948.21	14.42	17	.637

Model definitions. The Fully Sat. model is the fully saturated model of the observed data, which models the means and variances for both variables, and the phenotypic and cross-twin-cross-trait correlations between the two variables, separately for each twin in a pair and across zygosity. **5.** In the Bivariate model fitting, the constrained saturated model equates means, variances, phenotypic and cross-twin-cross-trait correlations across twins within a pair and across zygosity, for both variables of interest.

The best-fitting model (in bold) was the non-significant and most parsimonious model, as well as the one with the lowest AIC.

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference

Table S 8. ACE and ADE Bivariate twin model fit statistics and estimates between face orienting (proportion of first look to faces) and face

preference. Best-fitting model in bold.

Model	# Parameters	-2LL	df	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A.1	A.12	A.2	C/D.1	C/D.12	C/D.2	E.1	E.12	E.2
Fully Sat	32	1151.37	1040	-928.63	NA	NA	NA	NA	-	-	-	-	-	-	-	-	-
ACE	15	1167.82	1057	-946.18	Fully Sat.	16.45	17	.492	.15	.37	0	.04	.08	0	.81	.02	.53
ADE	15	1167.91	1057	-946.09	Fully Sat.	16.54	17	.486	< .01	.24	0	.21	.23	0	.79	.02	.52
AE	12	1169.33	1060	-950.67	ADE	1.43	3	.700	.19	.16	.29	-	-	-	.81	.02	.53
E	9	1211.82	1063	-914.18	ADE	43.92	6	< .001	-	-	-	-	-	-	1	.09	.91

Model definitions. The Fully Sat. model is the fully saturated model of the observed data, which models the means and variances for both variables, and the phenotypic and cross-twin-cross-trait correlations between the two variables, separately for each twin in a pair and across zygosity. The best-fitting model (in bold) was the non-significant and most parsimonious model, as well as the one with the lowest AIC.

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in $-2LL$ statistic between two models, distributed χ^2 .

Δdf = difference

A = additive genetic influences

C = shared environment influences

D = non-additive genetic influences

E = non-shared environment influences

A-C/D-E.1 = variance on Phenotype 1 (proportion of first look to face)

A-C/D-E.12 = variance on Phenotype 2 (face preference) that is shared with Phenotype 1 (proportion of first look to face)

A-C/D-E.2 = unique variance on Phenotype 2 (face preference)

Table S 9. Summary of the two GEE models including genome-wide polygenic scores (GPSs) for autism, ADHD, bipolar disorder, major depressive disorder, and schizophrenia, 10 principal component of ancestry, and age and sex, as predictors of face orienting (proportion of first look to faces) or face preference, with twin pair id as cluster-defining variable. Significant predictors are in bold (alpha threshold was adjusted for the number of outcomes, = .025).

Face orienting (% first look to face)	Beta	Uncorrected p -value
Age (in days)	0.00	.499
Sex	-0.03	.088
GPS autism	0.00	.716
GPS ADHD	-0.01	.176
GPS bipolar disorder	0.00	.942
GPS major depression disorder	0.02	.048
GPS schizophrenia	-0.02	.099
PC 2	-0.04	.979
PC 3	-1.87	.029

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

PC 4	2.56	.150
PC 5	3.58	.076
PC 6	-0.37	.848
PC 7	-0.68	.678
PC 8	-0.90	.578
PC 9	3.17	.097
PC 10	0.07	.968
PC 11	4.61	.017
Face preference (% on face)	Beta	Uncorrected <i>p</i>-value
Age (in days)	0.02	.003
Sex	-0.18	.058
GPS autism	-0.03	.601
GPS ADHD	-0.01	.872
GPS bipolar disorder	-0.07	.377
GPS major depression disorder	0.06	.289
GPS schizophrenia	0.07	.329
PC 2	-15.01	.041
PC 3	-1.15	.836
PC 4	-22.98	.053
PC 5	-7.18	.515
PC 6	-4.96	.620
PC 7	-10.87	.284
PC 8	2.42	.800
PC 9	20.66	.050
PC 10	0.56	.960
PC 11	19.02	.054

Appendix S1. Steps in gaze offline pre-processing:

- 1) Continuous raw eye tracking data was resampled to 60Hz
- 2) Off-screen gaze was marked as missing gaze
- 3) X and Y coordinates were averaged when binocular data was present (data from one eye was used when one eye was missing)
- 4) Large AOIs (centre, face, noise, car, bird, phone) were defined around each stimulus (see **Fig. S1**).
 - a. Raw data was assigned to AOIs, by coding logical vectors (n=5) of gaze samples inside (1) and outside (0) each AOI
 - b. AOI vectors were interpolated to fill in gaps of missing data shorter than 200ms (i.e. recode from 0 to 1)
 - c. Any runs of samples in an AOI vector with a length less than 50ms were recoded to 0 (trigger tolerance for AOI activation), to ensure that a minimum of 50ms of gaze data was accumulated in an AOI for a look to be computed.
- 5) For each trial and AOI we computed:
 - a. Whether the AOI was looked at, coded as true if at least a contiguous run of samples with 50 ms was identified inside the AOI.
 - b. The latency of the first look (if a. was true), coded as the duration from the start of the trial to the first sample in the AOI.
 - c. The looking ratio in the AOI, coded as the number of samples in AOI per number of samples in all AOIs.
- 6) For each trial we also computed:
 - a. The number of objects looked at, coded as the count of AOIs looked at (i.e., 5. a. was true).

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Figure S 1. Stimuli arrays of the face pop-out task with overlapping Areas of interest.

Figure S 2. Distribution and boxplots for 536 5-month-old infants' measures: proportion of trials where the infant looked first to each object at the start of the trial (first look) and proportion of time looking at each object throughout the trial, relative to time looking at all objects (ratio). Chance level highlighted as a dashed vertical line.

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Appendix S2. Sensitivity analyses for face orienting

To fulfil the pre-registered analyses plan, we report here the univariate analyses for face orienting operationalized as a composite average score between the proportion of trials that the infant looked at the face as the first AOI and the mean latency to look at the face (latencies shorter than 120 ms were excluded prior to averaging across valid trials) – see more details in the Methods section “Computation of primary measures”). The latency score was reversed and both measures were z-scored before averaging them. A higher score on this measure indicates a faster and larger face orienting. The twin correlations suggested genetic influences on the composite score ($ICC_{MZ}=.25$, CI: .08-.39; $ICC_{DZ}=.13$, CI: -.06-.30). The twin models confirmed these genetic effects but the assumptions for twin modelling were not met.

Table S 10. Descriptive statistics for face orienting composite score and latency to look at the face AOI.

	Overall	MZ Females	MZ Males	DZ Females	DZ Males	Skewness
N (twins)	536	135	158	116	127	
Face orienting composite	0 (0.81) -2.50-2.41	-0.05 (0.78) -2.14-1.88	-0.10 (0.83) -2.18-2.07	0.17 (0.84) -1.93-2.05	0.02 (0.79) -2.50-2.41	0.05
Latency (ms)	2.93 (1.82) 0.41-10.80	3.00 (1.73) 0.56-8.13	3.06 (1.93) 0.41-10.21	2.71(1.77) 0.50-9.01	2.90 (1.82) 0.54-10.80	1.06

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Table S 11. Saturated model for the composite score of face orienting (proportion of first look to face and latency), including covariates. Note this is not the same measure that it was reported in the manuscript (only proportion of first look to face).

Model	# Parameters	-2LL	df	AIC	Comparative fit with saturated model		
					$\Delta \chi^2$	Δdf	P value
Saturated	12	1484.49	524	436.49	-	-	-
Age	11	1494.07	525	444.07	9.59	1	.002
Sex	11	1486.12	525	436.12	1.64	1	.201
1.	10	1489.12	526	437.12	4.63	2	.099
2.	9	1493.34	527	439.34	8.86	3	.031
3.	7	1499.60	529	441.60	15.11	5	.010
4.	6	1499.61	530	439.61	15.13	6	.019

Model definitions. The baseline model is the fully saturated model of the observed data, which models the means and variances separately for each twin in a pair and across zygosity. **Age**. Testing the significance of age, **Sex**. Testing the significance of sex, **1.** Equating means across twins within a pair, **2.** Equating means across zygosity, **3.** Equating variances across twins within a pair, and **4.** Equating variances across zygosity (i.e., the constrained saturated model).

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

Table S 12. Univariate twin model fit statistics and parameter estimates for the composite score of face orienting. Note this is not the same measure that it was reported in the manuscript (proportion of first look to face). Best-fitting model in bold.

Model	# parameters	-2LL	df	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A	C	E
Fully Sat	12	1484.49	524	436.49	NA	-	-	-	-	-	-
ACE	6	1499.61	530	439.61	Fully Sat.	15.13	6	.019	.23	.02	.76

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

AE	5	1499.62	531	437.62	ACE	0.01	1	.942	.25 (.09-.39)	0	.75 (.61-.91)
CE	5	1500.49	531	438.49	ACE	0.87	1	.350	0	.19	.81
E	4	1509.38	532	445.38	ACE	9.76	2	.008	0	0	1

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

Table S 13. Bivariate twin model fit statistics for the composite score of face orienting and face preference. Best-fitting model in bold. Note this is not the same analyses that it was reported in the manuscript (which only included proportion of first look to face).

Model	# Parameters	-2LL	<i>df</i>	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A.1	A.12	A.2	C.1	C.12	C.2	E.1	E.12	E.2
Fully Sat	32	2845.27	1040	765.27	NA	NA	NA	NA	-	-	-	-	-	-	-	-	-
5.	15	2865.95	1057	751.95	Fully Sat	20.68	17	.241	-	-	-	-	-	-	-	-	-
ACE	15	2867.28	1057	753.28	Fully Sat.	22.01	17	.184	.22	.41	0	.04	.04	0	.75	.04	.51
ACE nested models																	
AE	12	2868.67	1060	748.67	ACE	1.40	3	.706	.25	.27	.19	-	-	-	.75	.04	.51
CE	12	2874.65	1060	754.65	ACE	7.37	3	.061	-	-	-	.19	.17	.18	.81	.07	.57
E	9	2912.25	1063	786.25	ACE	44.97	6	< .001	-	-	-	-	-	-	1	.18	.82

Model definitions. The Fully Sat. model is the fully saturated model of the observed data, which models the means and variances for both variables, and the phenotypic and cross-twin-cross-trait correlations between the two variables, separately for each twin in a pair and across zygosity. **5.** In the Bivariate model fitting, the constrained saturated model equates means, variances, phenotypic and cross-twin-cross-trait correlations across twins within a pair and across zygosity, for both variables of interest.

The best-fitting model (in bold) was the non-significant and most parsimonious model, as well as the one with the lowest AIC.

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Appendix S3. Sensitivity analyses for number of objects explored (0-20 seconds)

To fulfil the pre-registered analyses plan, we report here the univariate analyses for number of objects explored operationalized as the number of objects looked at in the whole trial (20 seconds, in contrast with the measure reported in the main manuscript where only the first half of the trial was included – see more details in the Methods section “Computation of primary measures”). As in the reported measure, the twin correlations suggested no familial influences on the number of objects explored in the entire trial ($ICC_{MZ}=.08$, CI: $-.10-.25$; $ICC_{DZ}=.12$, CI: $-.05-.28$). This shows that, the result of the variability in visual exploration being best explained solely by unique environmental factors (which include measurement error) does not seem to be driven by any ceiling effects of the distribution of this measure.

Table S 14. Descriptive statistics for the number of objects explored during the entire trial length. Note this is not the same measure that it is reported in the manuscript (number of objects explored in the first 0-10 seconds).

	Overall	MZ Females	MZ Males	DZ Females	DZ Males	Skewness
N (twins)	536	135	158	116	127	
# Objects explored (in 0-20 seconds)	4.21 (0.48) 2.50-5.00	4.19 (0.47) 2.67-5.00	4.25 (0.47) 2.83-5.00	4.19 (0.45) 2.67-5.00	4.21 (0.53) 2.50-5.00	-0.84

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Table S 15. Saturated model for the number of objects explored (0-20 seconds), including covariates. Note this is not the same measure that it is reported in the manuscript (number of objects explored in the first 0-10 seconds).

Model	# Parameters	-2LL	df	AIC	Comparative fit with saturated model		
					$\Delta \chi^2$	Δdf	P value
Saturated	12	722.93	524	-325.07	-	-	-
Age	11	723.35	525	-326.65	0.42	1	.515
Sex	11	724.00	525	-326.00	1.08	1	.300
1.	10	725.47	526	-326.53	2.54	2	.281
2.	9	725.86	527	-328.14	2.94	3	.402
3.	7	726.56	529	-331.44	3.63	5	.604
4.	6	727.13	530	-332.87	4.20	6	.650

Model definitions. The baseline model is the fully saturated model of the observed data, which models the means and variances separately for each twin in a pair and across zygosity. **Age.** Testing the significance of age, **Sex.** Testing the significance of sex, **1.** Equating means across twins within a pair, **2.** Equating means across zygosity, **3.** Equating variances across twins within a pair, and **4.** Equating variances across zygosity (i.e., the constrained saturated model). **5.** In the Bivariate model fitting, the constrained saturated model equates means, variances, phenotypic and cross-twin-cross-trait correlations across twins within a pair and across zygosity, for both variables of interest.

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models.

SI for the research paper *Infants' looking preferences for social versus non-social objects reflect genetic variation and are linked to later language development* by Portugal et al.

Table S 16. Univariate twin model fit statistics and parameter estimates for the number of objects explored (0-20 seconds). Note this is not the same measure that it is reported in the manuscript (number of objects explored in the first 0-10 seconds). Best-fitting model in bold.

Model	# Parameters	-2LL	df	AIC	Comparison model	$\Delta \chi^2$	Δdf	P value	A	C	E
Fully Sat	12	722.93	524	-325.07	NA	-	-	-	-	-	-
ACE	6	727.24	530	-332.76	Fully Sat.	4.32	6	.634	0	.10	.90
AE	5	727.82	531	-334.18	ACE	0.57	1	.449	.11	0	.89
CE	5	727.24	531	-334.76	ACE	0.00	1	> .999	0	.10 (0-.22)	.90 (.78-1)
E	4	729.80	532	-334.20	ACE	2.55	2	.279	0	0	1

-2LL = fit statistic, which is minus two times the log-likelihood of the data.

df = degrees of freedom

AIC, an alternative fit index. Lower values denote better model fits.

$\Delta \chi^2$ = difference in -2LL statistic between two models, distributed χ^2 .

Δdf = difference in degrees of freedom between two models