

## BIROn - Birkbeck Institutional Research Online

Portugal, A.M. and Taylor, M.J. and Viktorsson, C. and Nystrom, P. and Li, D. and Tammimies, K. and Ronald, Angelica and Falck-Ytter, T. (2022) Pupil size and pupillary light reflex in early infancy: heritability and link to genetic liability to schizophrenia. *Journal of Child Psychology and Psychiatry* 63 (9), pp. 1068-1077. ISSN 0021-9630.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/49717/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html> or alternatively contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).

# Pupil size and pupillary light reflex in early infancy: heritability and link to genetic liability to schizophrenia

Ana Maria Portugal,<sup>1</sup>  Mark J. Taylor,<sup>2</sup>  Charlotte Viktorsson,<sup>3</sup> Pär Nyström,<sup>4</sup>  
Danyang Li,<sup>1,5</sup> Kristiina Tammimies,<sup>1,5</sup>  Angelica Ronald,<sup>6</sup> and Terje Falck-Ytter<sup>1,3,7</sup>

<sup>1</sup>Department of Women's and Children's Health, Center of Neurodevelopmental Disorders (KIND), Centre for Psychiatry Research, Karolinska Institutet & Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden; <sup>2</sup>Department of Medical Epidemiology & Biostatistics, Karolinska Institutet, Stockholm, Sweden; <sup>3</sup>Development and Neurodiversity Lab, Department of Psychology, Uppsala University, Uppsala, Sweden; <sup>4</sup>Uppsala Child & BabyLab, Department of Psychology, Uppsala University, Uppsala, Sweden; <sup>5</sup>Astrid Lindgren Children's Hospital, Karolinska University Hospital, Region Stockholm, Stockholm, Sweden; <sup>6</sup>Department of Psychological Sciences, Birkbeck, University of London, London, UK; <sup>7</sup>Swedish Collegium for Advanced Study, Uppsala, Sweden

**Background:** Measures based on pupillometry, such as the pupillary light reflex (PLR) and baseline pupil size, reflect physiological responses linked to specific neural circuits that have been implicated as atypical in some psychiatric and neurodevelopmental conditions. **Methods:** We investigated the contribution of genetic and environmental factors to the baseline pupil size and the PLR in 510 infant twins assessed at 5 months of age (281 monozygotic and 229 dizygotic pairs), and its associations with common genetic variants associated with neurodevelopmental (autism spectrum disorder and attention deficit hyperactivity disorder) and mental health (bipolar disorder, major depressive disorder and schizophrenia) conditions using genome-wide polygenic scores (GPSs). **Results:** Univariate twin modelling showed high heritability at 5 months for both pupil size ( $h^2 = .64$ ) and constriction in response to light ( $h^2 = .62$ ), and bivariate twin modeling indicated substantial independence between the genetic factors influencing each ( $r_G = .38$ ). A statistically significant positive association between infant tonic pupil size and the GPS for schizophrenia was found ( $\beta = .15$ ,  $p = .024$ ), while there was no significant association with the GPS for autism or any other GPSs. **Conclusions:** This study shows that some pupil measures are highly heritable in early infancy, although substantially independent in their genetic etiologies, and associated with common genetic variants linked to schizophrenia. It illustrates how genetically informed studies of infants may help us understand early physiological responses associated with psychiatric disorders which emerge much later in life. **Keywords:** Pupillometry; pupillary light reflex; infancy; twin design; polygenic risk scores; schizophrenia.

## Introduction

The pupillary light reflex (PLR) is a physiological response to retinal or attended stimulus luminance serving to regulate the amount of light reaching the retina. Individual variation in the PLR (the relative amplitude the pupil constricts or the latency to initiate the constriction response) is thought to reflect altered sensory reactivity, and the measure has a wide range of clinical applications, including detection of parasympathetic dysfunction (Wang et al., 2016) and cholinergic deficits (Hall & Chilcott, 2018). Although an involuntary reaction, it has been shown to be modulated by top-down cognitive processes (Ebitz & Moore, 2017; Mathôt, Linden, Grainger, & Vitu, 2015). While the sudden constriction to light seen in the PLR is linked primarily to the parasympathetic nervous system and draws heavily on cholinergic synaptic transmission (Hall & Chilcott, 2018), fluctuations in the pupil more generally reflect also other brain systems, including the sympathetic activity and the locus coeruleus–norepinephrine (LC-NE) system more specifically

(Aston-Jones & Cohen, 2005; Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Murphy, Robertson, Balsters, & O'Connell, 2011), which index ongoing psychological processes related to arousal and attention (Laeng, Sirois, & Gredebäck, 2012). Tonic levels (baseline levels) of pupil dilation thus both reflect the amount of light reaching the eye, and the level of attention and arousal of the individual in that context.

In recent years, there has been an increasing interest in the link between the PLR early in life and autism spectrum disorder (ASD). In children, adolescents, and adults, a hypo-sensitive response (slower latency and reduced relative amplitude) has been shown at a group level (Daluwatte et al., 2013; Fan, Miles, Takahashi, & Yao, 2009) and associated with ASD traits (DiCriscio & Troiani, 2017). In contrast, early in life, a hypersensitive PLR has been seen in infants with a later ASD diagnosis and associated with symptom severity (Nyström et al., 2018; Nyström, Gredebäck, Bölte, & Falck-Ytter, 2015). Atypical PLR responses have also been described in psychiatric conditions such as schizophrenia, where a hypo-sensitive response has been found at a group level (Hakarem, Lidsky, &

Conflict of interest statement: No conflicts declared.

Sutton, 1966) and in association with more negative symptoms (i.e. symptoms related to withdrawal, lack of motivation and interest were associated with a slower latency to constrict; Fattal et al., 2020).

Twin studies help us understand the nature of individual differences in terms of genetic and environmental contributions to variability in traits. So far, no study that we are aware of investigated pupil measures such as the PLR in infants in a genetically informed design. This is striking given the link between the PLR and strongly heritable neurodevelopmental conditions like ASD and schizophrenia (Daluwatte et al., 2013; Fan et al., 2009; Fattal et al., 2020). The current infant twin study aimed to clarify the etiological background and genetic links of the PLR (the relative constriction amplitude and the latency) and the pupil size (baseline pupil diameter) at 5 months of age. The primary aim was to quantify the contribution of genetic and environmental factors to the PLR and baseline pupil size, and to the covariance between the different pupil measures. In light of previous research indicating significant genetic effects on early brain development (e.g. morphology measures; Maggioni, Squarcina, Dusi, Diwadkar, & Brambilla, 2020), we expected all pupil measures to be heritable. Our second aim was to utilize genome-wide polygenic scores (GPSs, a genotype-based individual score that summarizes the estimated effect of many trait-associated alleles) to assess whether the common genetic architecture known to be associated with psychiatric and neurodevelopmental conditions, particularly ASD and schizophrenia, associates with infant pupil parameters.

## Methods and materials

Participants were part of the BabyTwins Study Sweden (BATSS), a longitudinal infant twin study (Falck-Ytter et al., 2021). BATSS has a classic twin design, in which similarity is compared between monozygotic twin pairs (MZ; who share 100% of their segregating alleles) and dizygotic same-sex pairs (DZ; who on average share 50% of their segregating alleles). Twins were identified via the Swedish population registry and contacted by letter. Only the greater Stockholm area was targeted due to the need to come in person to the lab. Inclusion and exclusion criteria, checked via telephone interview, were opposite-sex twin pairs, diagnosis of epilepsy, known presence of a genetic syndrome, vision or hearing impairments, very premature birth (prior to week 34), presence of developmental or medical conditions likely to affect brain development (e.g. Cerebral Palsy), and infants where none of the biological parents were involved in the infant's care. Parents also needed to be willing to disclose information about family medical and psychiatric history, demographics, and birth delivery. From 2016 to 2020, 311 families (29% of the entire population of same-sex twins born in the area; 57% MZ pairs, as estimated based on DNA sampled from all infants) took part in an in-person multi-methods assessment at 5 months (testing site at Karolinska Institutet). The study sample has a high socioeconomic status, and it comprises of mostly Swedish-born families (90% of twin pairs had at least one parent born in Sweden). Recruitment summary and sample demographics are fully reported in the BATSS description paper (Falck-Ytter

et al., 2021). Parents gave informed consent to take part. BATSS was approved by the regional ethics board in Stockholm and was conducted in accordance with the Declaration of Helsinki.

For the current analyses, 28 twins were excluded due to parent-reported twin-to-twin transfusion syndrome (12 pairs), report of seizures at the time of birth (2 individuals from different pairs), very low birth weight (<1.5 kg, 1 individual), and report of spina bifida (1 individual). In addition, 23 infants did not take part in the eye-tracking assessment due to technical reasons (3 pairs), lack of time (1 pair plus 1 individual), bad calibration (1 individual), and tiredness (4 pairs plus 5 individuals). From the infants that took part in the assessment, 61 infants (9 pairs plus 43 individuals) did not have enough valid trials to yield pupil measures (see below criteria for measurements), so these were excluded from the analysis. The final sample was of 510 individuals (280 incomplete pairs, i.e. pairs with at least one twin with data). We found no statistically significant differences between the excluded and included infants in terms of either parental education level, family income, sex, age or any of the GPSs used in this study.

## Measures

The PLR was measured during a lab visit at 5 months of age which included other paradigms not related to the current study (Falck-Ytter et al., 2021). During this visit, both twins performed the PLR task; one twin was assessed at a time.

### Eye tracking protocol

To record gaze and pupil data a Tobii TX300 eye-tracker was used (sampling rate of 120 Hz) with scripts written for the Eurosibs study (Jones et al., 2019) in MATLAB (version R2013b, MathWorks, Natick, MA, USA) and Psychtoolbox (version 3.0.12). An initial 5-point calibration was called before the start of the task battery, which lasted for about 10 min and involved rotations of free-viewing of dynamic and static scenes (mixture of social and nonsocial content), gaze-contingent gap-overlap trials, PLR measurements, and postcalibration sequences. Here, we analyzed only pupil data collected during the presentation of 12 trials that elicited a PLR. These 12 PLR measurements were not collected in sequence, rather they were always interspersed with other stimuli with varying luminosity.

### The pupillary light reflex task

For each trial, lasting approximately 6 s, the stimulus consisted of a small central fixation point on a black background (0.9 lux) that flashed white (190 lux) for 75 ms with a random onset between 1,600 and 2,400 ms. In line with previous research, the PLR was evaluated both in terms of its relative constriction and its latency. The relative pupil constriction was calculated as in Fan et al. (2009) by the formula  $(A_0^2 - A_m^2) / A_0^2$ , where  $A_0$  is the average pupil diameter before the onset of the PLR (during an

interval starting 100 ms before and ending at the PLR onset, as determined by the PLR latency), that is, the mean baseline, and  $A_m$  is the minimum pupil diameter in the interval 500–1,500 ms relative to the stimulus onset (Fan et al., 2009). Following previous work (Nyström et al., 2015, 2018), the PLR latency was defined at the acceleration minimum between flash onset and minimum pupil diameter. Computational processing was done in MATLAB using the TimeStudio framework (Nyström, Falck-Ytter, & Gredebäck, 2016) and is described elsewhere (Nyström et al., 2015, 2018). All trials were inspected by a researcher (blind to zygosity). Invalid trials (with apparent artifacts and implausible responses) for each eye were first automatically marked, and then visually rejected. Only one eye per trial was kept and infants with less than three valid trials were excluded. Mean values from valid trials were computed to get the outcome variables (relative constriction; latency of the response; and baseline pupil diameter). Short-term internal stability of the pupil measures was estimated by intra-class correlations (two-way random effects) using mean values of the first six trials (test block) and of the last six trials (retest block) – relative constriction had moderate reliability ( $ICC_{\text{twin } 1} = .60/ICC_{\text{twin } 2} = .58$ ), baseline pupil diameter had good reliability ( $ICC = .83/.87$ ), whereas latency of the response had poor reliability ( $ICC = .12/.18$ ).

### Genome-wide polygenic scores

The DNA samples were genotyped using Infinium Global Screening Array (Illumina, San Diego, CA, USA). Genotype quality control and processing was done using standard procedures and described elsewhere (Falck-Ytter et al., 2021). GPSs were calculated based on the most recent and largest genome-wide association studies for a range of neurodevelopmental and psychiatric conditions [ADHD (Demontis et al., 2019), ASD (Grove et al., 2019), bipolar disorder (Stahl et al., 2019), major depressive disorder (Howard et al., 2019), and schizophrenia (Consortium, Ripke, Walters, & O'Donovan, 2020)]. The GPSs were calculated using the PRS-CS (polygenic prediction via Bayesian regression and continuous shrinkage priors) method (Ge, Chen, Ni, Feng, & Smoller, 2019). For the GPS analyses, the first 10 principal components of ancestry were included as covariates in the models. In the sample, only one pair was of completely non-European genetic ancestry. This pair was retained in the analysis since their GPSs for all five conditions above-mentioned were below 3 standard deviations of the sample.

### Data analysis

The analysis plan for this study was preregistered in OSF (<https://osf.io/5kzh2>), after data collection,

cleaning, and extraction, but prior to any analysis of the data. PLR latency, PLR relative constriction, and pupil baseline had skewness values between  $-1$  and  $1$  (.04 for Latency, .23 for Relative Constriction and .59 for Baseline) so they were not transformed. All measures were standardized. R software (version 4.0.0) was used for all analyses.

For twin models, the OpenMx package (version 2.18.1; Neale et al., 2016) with full-information maximum likelihood estimation was used. Even though these methods allow for partially complete pairs (one twin missing data) to be included, they were not included in the current analysis because the twins with missing data were also missing certain covariates (which have to be complete for all cases).

The effects of covariates on the PLR parameters were estimated within the twin model. These included the number of valid trials (a proxy of data quality), the year of data collection (a proxy of possible changes in luminance levels of the testing room, even though no systematic differences were detected), age (in days, same for all twins, except for six pairs where the average age was calculated), and sex. In a deviation from the preregistered plan (covariates would be added if found to be associated with the relevant outcome PLR measures), these variables were added as covariates in all twin analyses; the findings do not change between approaches.

Univariate twin models estimate the relative contribution of genetic and environmental factors to the variation in a phenotype. This variation can be partitioned into additive genetic influence (A, heritability, which increases twin similarity), nonshared 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 nonadditive genetic (D) effects (D and C variance cannot be estimated simultaneously with twin data alone, because they confound one another). These components can be estimated by comparing the resemblance in MZ and DZ pairs. When MZ pairs are more similar than DZ pairs we can infer that variance in a phenotype is influenced by genetic factors; when DZ pairs are more than half as similar as MZ pairs, we can infer that shared environment influences the variance in that phenotype.

Univariate twin models were fit separately for each pupil measure. First, to test for the assumptions of equality of mean and variances and to derive twin correlations, a constrained saturated model, in which the means and variances were constrained to be equal across twin order and zygosity, was fit. Next, using either a model with A, C, and E (ACE model, where A stands for additive genetic effects, C for shared environmental effects, and E for

nonshared environmental effects) or a model with A, D, and E (ADE model, where A stands for additive genetic effects, D for nonadditive or dominance genetic effects, and E for nonshared environmental effects), we estimated and tested the genetic and environmental contributions to the variance in the measures. The best-fitting model was the most parsimonious (fewer number of parameters estimated) and nonsignificant model (meaning that there was no decrement in fit compared to the genetic model, indexed by the  $\chi^2$  distribution).

Bivariate twin modeling allows the relative contributions of genetic and environmental factors to the covariation between two measures to be estimated. It assumes that the observed correlation between measures can be decomposed into genetic, shared environmental, and unique environmental effects. In the bivariate case, cross-trait cross-twin correlations (CTCTs) are compared, that is, the correlation between one measure for one twin and another measure for their co-twin, between MZ and DZ pairs.

Bivariate twin models were fitted between phenotypes that were phenotypically correlated and shown to be heritable in univariate analyses. To test for the assumption of the equality of phenotypic and CTCTs in the data and to derive these correlations, a constrained saturated model, in which the phenotypic and CTCTs were constrained to be equal across twin order and zygosity (for the phenotypic correlation), was fitted. Next, using a bivariate ACE or ADE model and a correlated factors solution, we examined genetic and environmental correlations and bivariate heritability (i.e. the proportion of the phenotypic correlation that is explained by genetic factors).

To investigate the associations between the pupil measures and GPSs, a generalized estimating equation (GEE) framework (using the *drgee* package; Zetterqvist & Sjölander, 2015), with cluster-robust SE to account for family relatedness, was used to derive linear regression Beta estimates and *p*-values. Pupil outcomes were regressed on age and sex before analysis. *P*-thresholds were adjusted for the number of outcomes. To estimate effect sizes ( $\Delta R^2$ ) for model comparison, the  $R^2$  of corresponding models were derived and compared to the  $R^2$  of the null model (i.e. the model with only covariates included).

Associations with the Infant and Toddler Sensory Profile, which were planned in the preregistration, are reported in Appendix S1 (no statistically significant associations were found, Tables S1 and S2).

## Results

### Twin modeling analyses

Sample descriptive statistics are presented in Table 1. In terms of the effects of the covariates, there were no significant covariate effects for latency to constrict (see Table S3 in Appendix S2). Age was

positively related to relative constriction ( $\beta = .02$ ,  $p < .001$ , see Table S4) and to baseline ( $\beta = .01$ ,  $p = .025$ , see Table S5), that is, older infants had increased constriction amplitudes and bigger baseline pupils. Sex was related to baseline diameter ( $\beta = .37$ ,  $p = .001$ , infant boys had, on average, .16 mm larger pupils, see Table S5). These patterns of age and sex effects are in line with previous literature with infants and toddlers (Kercher et al., 2020). Number of valid trials and year of testing were statistically associated with baseline diameter (more valid trials yielded a smaller pupil size,  $\beta = -.04$ , and pupil size was smaller in the last year of data collection,  $\beta = -.14$ ; both  $p < .01$ , see Table S5).

Table 2 shows the twin correlations for the pupil parameters. Higher MZ than DZ correlations suggests heritable influences on baseline and relative constriction. Low shared environmental influences were suggested because the DZ correlations were only marginally greater than half the MZ correlations. For latency, there was very low similarity within both MZ and DZ pairs, suggesting no familial effects on the phenotype. For this reason, further results for latency were not reported (but see Appendix S2).

Univariate twin modeling (see model assumptions tests in Tables S4–S6) confirmed the different patterns of genetic and environmental contributions to the pupil measures. Table 3 shows the twin models fit statistics and the standardized estimates.

For relative constriction, there were significant high genetic effects (best-fitting AE model  $A = .62$ , CI: .51–.70); while the contribution of the shared environment was nonsignificant (ACE model  $C = .12$ , CI: 0–.42). For baseline diameter, high genetic effects were significant (best-fitting AE model  $A = .64$ , CI: .54–.72), and the shared environment had a minimal and nonsignificant contribution (ACE model  $C = .03$ , CI: 0–.37).

The phenotypic correlation between baseline and relative constriction was moderate ( $r_{Ph} = .38$ , CI: .28–.46), suggesting that infants with a larger pupil at baseline had higher constriction amplitudes. A saturated model (see the bivariate model assumptions tests in Table S7) was used to derive the CTCT correlations, which suggested genetic influences on the phenotypic correlation because the MZ CTCT ( $r_{CTCT\ MZ} = .24$ , CI: .13–.34) was higher than the DZ CTCT ( $r_{CTCT\ DZ} = .06$ , CI:  $-.09$ –.21). Because the correlation in DZ was less than half of the correlation in MZ twins, some nonadditive genetic effects could explain the phenotypic correlation between the measures. However, given the sample size and the small (although nonsignificant) contribution of shared environment seen in the univariate models, we did not fit an ADE model. Table 4 shows the model fit statistics for the bivariate model. An AE model fitted the data best, so this model was used to derive separately the genetic correlations between pupil baseline and relative constriction (Figure 1).

**Table 1** Descriptive statistics

	Overall	MZ females	MZ males	DZ females	DZ males
<i>N</i> (twins)	510	129	152	116	113
Age (days)	167.70/8.59 [145–203]	168.01/9.16 [153–194]	167.41/8.09 [150–187]	168.03/8.24 [153–189]	167.41/8.96 [145–203]
Pupillary light reflex					
Number of valid trials	8.48/2.85 [3–12]	8.33/2.67 [3–12]	8.47/2.92 [3–12]	8.68/2.79 [3–12]	8.44/3.04 [3–12]
Mean latency (s)	0.35/0.04 [0.18–0.52]	0.35/0.05 [0.24–0.49]	0.36/0.04 [0.24–0.52]	0.35/0.05 [0.22–0.49]	0.35/0.04 [0.18–0.45]
Mean relative constriction (%)	28.78/7.75 [2.42–51.90]	28.40/7.63 [11.61–51.90]	29.03/7.67 [9.90–48.57]	28.54/7.98 [11.90–50.46]	29.11/7.85 [2.42–47.53]
Mean baseline (mm)	3.23/0.46 [2.27–4.96]	3.17/0.48 [2.31–4.96]	3.33/0.48 [2.27–4.85]	3.13/0.38 [2.27–4.19]	3.28/0.45 [2.39–4.69]

Statistics presented as mean/*SD* [min max]. DZ, dizygotic; MZ, monozygotic.

**Table 2** Twin correlations for the PLR measures, separate for MZ and DZ pairs. 95% confidence intervals are shown in parentheses

	N (Twin pairs)	Latency	Relative constriction	Baseline
MZ	127	.15 [–.03–.32]	.61 [.49–.70]	.64 [.54–.72]
DZ	103	.28 [.10–.44]	.36 [.19–.51]	.34 [.12–.51]

The proportion of the phenotypic correlation that was explained by genetic influences was .63 (CI: .41–.79; i.e. the bivariate heritability) and the remainder was explained by unique environment (.37, CI: .21–.59). Genetic correlations between the two measures indicated moderate overlap in the genetic factors underlying the measures ( $r_A = .38$ , CI: .23–.52). A similar result was seen for the unique environmental correlation ( $r_E = .38$ , CI: .23–.51) indicating that the unique environmental influences partially overlapped.

### Associations with common genetic variants underlying neurodevelopmental and psychiatric conditions

Preregistered exploratory GEE analyses were conducted to test the associations between pupil size and PLR constriction with a set of neurodevelopmental and psychiatric GPSs (ASD, ADHD, bipolar disorder, major depressive disorder, and schizophrenia) in two single models.

For pupil size at baseline, the GPS for schizophrenia was found to be positively associated,  $\beta = .15$ ,  $p = .024$  (alpha threshold adjusted for the number of pupil outcomes is .025),  $\Delta R^2 = .016$ , when controlling for the other neurodevelopmental and psychiatric GPSs, see Table 5. The GPS for schizophrenia was still found to be marginally associated with the baseline when the schizophrenia GPS was included in a model without any other GPSs,  $\beta = .14$ ,  $p = .031$ ,  $\Delta R^2 = .012$ . There were no statistically significant associations between the GPS for schizophrenia and the study covariates (age, sex,

**Table 3** Univariate twin model fit statistics and parameter estimates. Best-fitting model in bold

Model	# Parameters	–2LL	<i>df</i>	AIC	Comparison model	$\Delta\chi^2$	$\Delta df$	<i>p</i> Value	A	C	E
Relative constriction											
Fully Sat	14	1198.18	446	306.18	NA	–	–	–	–	–	–
ACE	8	1200.17	452	296.17	Fully Sat.	1.99	6	.920	.49	.12	.39
<b>AE</b>	<b>7</b>	<b>1200.64</b>	<b>453</b>	<b>294.64</b>	<b>ACE</b>	<b>0.47</b>	<b>1</b>	<b>.491</b>	<b>.62</b>	<b>0</b>	<b>.39</b>
CE	7	1207.66	453	301.66	ACE	7.49	1	.006	0	.49	.51
E	6	1271.98	454	363.98	ACE	71.81	2	<.001	0	0	1
Baseline											
Fully Sat	14	1155.17	446	263.17	NA	–	–	–	–	–	–
ACE	8	1163.45	452	259.45	Fully Sat.	8.28	6	.228	.61	.03	.36
<b>AE</b>	<b>7</b>	<b>1163.47</b>	<b>453</b>	<b>257.47</b>	<b>ACE</b>	<b>0.02</b>	<b>1</b>	<b>.884</b>	<b>.64</b>	<b>0</b>	<b>.36</b>
CE	7	1174.48	453	268.48	ACE	11.03	1	.001	0	.54	.46
E	6	1249.67	454	341.67	ACE	86.22	2	<.001	0	0	1

Model definitions. The Fully Sat. 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.

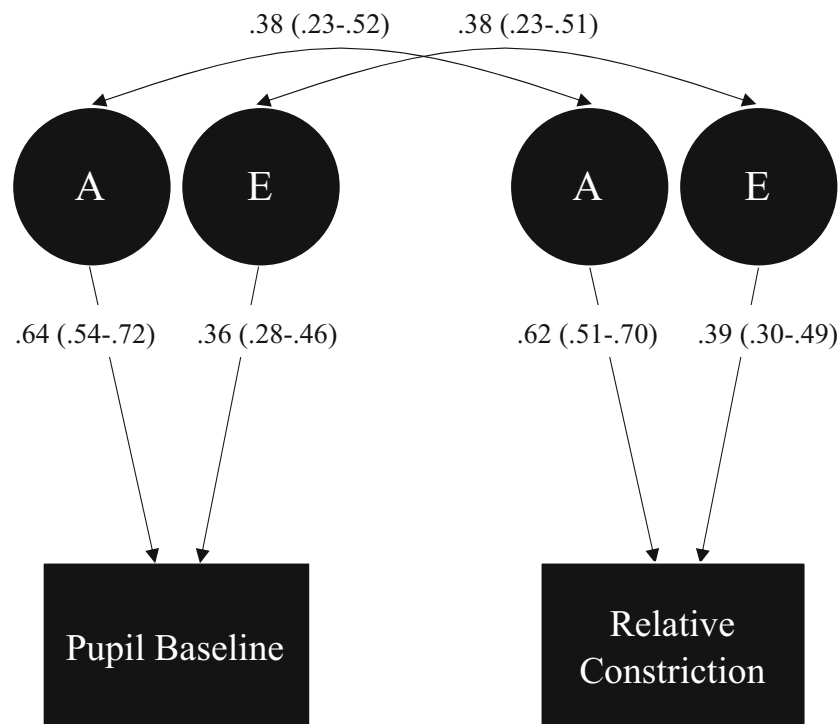
–2LL = fit statistic, which is minus two times the log-likelihood of the data; *df* = degrees of freedom; AIC, an alternative fit index, lower values denote better model fits;  $\Delta\chi^2$  = difference in –2LL statistic between two models, distributed  $\chi^2$ ;  $\Delta df$  = difference in degrees of freedom between two models.

**Table 4** Bivariate twin model fit statistics. Best-fitting model in bold

Model	# Parameters	-2LL	df	AIC	Comparison model	$\Delta\chi^2$	$\Delta df$	p Value	rA	rC	rE
Fully Sat	36	2270.68	884	502.68	NA	NA	NA	NA	-	-	-
ACE	19	2289.04	901	487.04	Fully Sat.	18.36	17	.367	.64	-1	.36
<b>AE</b>	<b>16</b>	<b>2292.15</b>	<b>904</b>	<b>484.15</b>	<b>ACE</b>	<b>3.11</b>	<b>3</b>	<b>.375</b>	<b>.38</b>	-	<b>.38</b>
CE	16	2305.04	904	497.04	ACE	16.00	3	.001	-	.34	.42
E	13	2450.21	907	636.21	ACE	161.17	6	<.001	-	-	.38

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 nonsignificant and most parsimonious 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  $\chi^2$ ;  $\Delta df$  = difference.



**Figure 1** The AE bivariate model was the best fit for baseline and relative constriction. This was used to derive the genetic correlations between phenotypes using a correlated factors solution

year of data collection, and number of valid trials). No other statistically significant associations between pupil measures and GPSs were found (see Table S8).

A preregistered hypothesis-driven GEE model was also computed to predict PLR relative constriction using the GPS for ASD alone, but no statistically significant association was found ( $\beta = .01$ ,  $p > .25$ , see Table S9).

## Discussion

This study is the first, to our knowledge, to quantify the influence of genetic and environmental factors on pupil size and the PLR using a classic twin design, and to report on the associations of these measures with common genetic variants for neurodevelopmental and psychiatric conditions. Due to the recent interest in these measures as potential antecedent

markers of neurodevelopmental conditions (Hall & Chilcott, 2018; Kercher et al., 2020), we focused on the early infancy period. As hypothesized, for relative constriction amplitude and pupil size at baseline, our results suggested strong genetic effects. The bivariate analysis revealed that relative constriction amplitude and pupil baseline diameter correlated modestly at the phenotypic level and that although genetic factors played a substantial role in the observed overlap between the two measures, most of the etiological influences on these measures were unique to each measure (genetic and unique environment correlations were moderate). This suggests that there are some distinct biological and environmental factors underlying the pupil baseline size versus the relative constriction amplitude. Distinct genetic influences on these measures are not unexpected given that the pupil baseline indexes tonic activity in the parasympathetic and sympathetic

**Table 5** Summary of the GEE model including genome-wide polygenic scores (GPSs) for ASD, ADHD, bipolar disorder, major depressive disorder, and schizophrenia and 10 principal component of ancestry as predictors of the pupil baseline size (regressed on age and sex) and twin pair id as cluster-defining variable. Significant predictors are in bold (*p*-threshold for significance is .025)

	Beta estimate	<i>p</i> -Value (uncorrected)
GPS ASD	0.00	.960
GPS ADHD	−0.01	.880
GPS bipolar disorder	−0.09	.250
GPS major depression disorder	0.03	.552
<b>GPS schizophrenia</b>	<b>0.15</b>	<b>.024</b>
PC 2	1.10	.904
PC 3	−2.41	.659
PC 4	10.93	.362
<b>PC 5</b>	<b>28.05</b>	<b>.021</b>
PC 6	12.83	.244
PC 7	14.63	.146
PC 8	0.14	.988
PC 9	−6.96	.496
PC 10	20.95	.060
PC 11	6.05	.630

nervous system [correlated with the LC-NE system activity system (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Murphy et al., 2011)], whereas the constriction response is a phasic measure linked to the parasympathetic system and cholinergic transmission more specifically (Hall & Chilcott, 2018). The observed partly overlapping unique environmental influences may be explained by the fact that these phenotypes are derived from the same trials; hence biomechanical factors and top-down modulation might contribute equally (also, although minimized by trial rejection, equipment failure, lightning conditions, child movement or crying might add covariance between measures). In addition, there may be early life unique experiences (such as early illnesses) playing a role in individual differences in pupil measures.

Regarding associations with molecular genetic GPS, our results suggested a small positive association between baseline pupil size and genetic risk for schizophrenia (more genetic predisposition was associated with a larger baseline pupil diameter), but do not add support to the previously reported link between the PLR relative constriction and risk for ASD at this age. The pupil size at the baseline is a measure of the average pupil diameter just before the onset of the PLR (in the 100 ms before the start of the constriction response), recorded while the infant was presented with an animated central fixation stimulus, but likely reflecting general arousal and engagement with the experiment as well (different stimuli, including visual arrays and videoclips, preceded the trials). The association with schizophrenia GPS is an interesting finding given

the literature suggesting that this tonic pupil size is correlated with the locus coeruleus–norepinephrine (LC-NE) system activity (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Murphy et al., 2011) and that LC-NE is implicated in a range of processes that are related to psychopathological development (e.g. arousal regulation, effort allocation, uncertainty processing, etc.). Although norepinephrine has been linked to schizophrenia (Mäki-Marttunen, Andreassen, & Espeseth, 2020; Maletic, Eramo, Gwin, Offord, & Duffy, 2017), it is not a neurotransmitter that is of current primary interest in the condition. However, given the general role the LC-NE plays for cortical function and development, a link with schizophrenia vulnerability is plausible. In any case, this finding was driven by exploratory analysis and requires replication, which we expect can be readily done using existing samples providing they have pupillometry and genotyped data.

For latency to constrict the pupil, twin similarity was small, with twin analysis failing to show any genetic or shared environmental effects. These results are indicative that only unique environment, which includes measurement error (a known factor affecting infant samples employing eye-tracking), explained the variance in latency to constrict, which together with the estimated poor internal stability, suggests this measure has low reliability at this age. For this reason, the hypothesized link between schizophrenia and PLR latency was not investigated. Overall, this finding argues for more stringent quality control in developmental studies using latency measures.

This study has some limitations. For both pupil size and relative constriction, shared environmental effects were nonsignificant, however, we acknowledge that the relatively small sample can be underpowered to detect these influences. This probably limited the estimation of the bivariate parameters as well, some of which had wide confidence intervals. While our sample included 29% of the same-sex twin population from the target area, it had a higher average SES (indexed by maternal formal education and family income) considering Stockholm normatives (which are already high compared to global standards) (Falck-Ytter et al., 2021). Our results need to be considered in light of these sample characteristics, as heritability estimates may vary in samples where SES has a broader distribution (e.g. Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003).

An important current limitation for studies assessing associations between phenotypes and GPSs is that the predictive power of a GPS is dependent on the original GWA study that the GPS calculation is based on (genome-wide association study used to identify the genetic variants associated with a trait). At the time of the present study, the available GWAS



for ASD was from Grove and colleagues (2019, discovery  $n = 46,350$  and SNP heritability = .19). In comparison, the most powerful GPS of a psychiatric condition is for schizophrenia, from the Schizophrenia Working Group of the Psychiatric Genomics Consortium (2020, discovery  $n = 105,318$  and SNP heritability = .41). It is the GPS for schizophrenia that we found preliminary evidence of being associated with baseline pupil size. The comparatively high power for the GPS in schizophrenia relative to the other neurodevelopmental/psychiatric phenotypes hinders strong conclusions about the specificity of this genetic association to pupil baseline and urges better powered GWAS studies for ASD to become available. Further, GWAS-driven GPSs only capture common genetic variants associated with a condition, but it may be that other genetic variants associated with ASD (e.g. rare variants) are associated with the pupil phenotypes, but this was out of the scope of this paper. Alternatively, PLR atypicalities associated with ASD may be better captured in other developmental windows than the one captured by our current study, as it has been reported that these vary considerably over the first two years of life (Fish et al., 2021).

## Conclusions

This is the first study reporting on the relative contribution of genetic and environmental influences to individual differences in tonic pupil size and the PLR and their associations with common genetic variants linked to neurodevelopmental and psychiatric conditions. Our results demonstrate that already in early infancy, the pupil size and the relative constriction in response to light are under substantial genetic influence, but that the genetic factors for these measures are to a large extent nonoverlapping. This supports previous research indicating distinctness in terms of biological mechanisms and low phenotypic association between the measures. Our results also provide evidence of an association between pupil size in infancy and genetic predisposition to schizophrenia. The study illustrates how genetically informed designs combined with early life psychophysiological measures can improve our understanding of human brain development and its disorders.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

**Appendix S1.** Associations between the pupil measures and the Infant/Toddler Sensory Profile.

**Appendix S2.** Twin modeling of latency to constrict.

**Table S1.** Descriptive statistics.

**Table S2.** Beta estimated from the GEE model of hyposensitivity and hypersensitivity scales of the ITSP predicting the PLR response.

**Table S3.** Univariate Twin Model Fit statistics and parameter estimates for the PLR latency to constrict measure.

**Table S4.** Saturated model for PLR relative constriction including covariates.

**Table S5.** Saturated model for PLR baseline including covariates.

**Table S6.** Saturated model for PLR Latency including covariates.

**Table S7.** Assumption testing for the bivariate model between PLR Relative Constriction and the Baseline.

**Table S8.** Summary of the GEE model including genome-wide polygenic scores (GPSs) for ASD, ADHD, bipolar disorder, major depressive disorder, and schizophrenia and 10 principal component of ancestry as predictors of the PLR relative constriction (regressed on age and sex) and twin pair id as cluster-defining variable.

**Table S9.** Summary of the pre-registered GEE model including genome-wide polygenic score (GPSs) for ASD, and 10 principal component of ancestry as predictors of the PLR relative constriction (regressed on age and sex) and twin pair id as cluster-defining variable.

## Acknowledgements

The authors thank all participating families, as well as research assistants Linnea Hamrefors, Joy Hättesstrand, Lynnea, Myers, Johanna Kronqvist, Sofia Jönsson, Anna Kernell, Carolin Schreiner, Sophie Lingö, Angelinn Liljebäck, Isabelle Enedahl, Matthis Andreasson, Lisa Belfrage, Mattias Savallampi, Isabelle Ocklind and Hjalmar Nobel Norrman. 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 (Initial Training Network BRAINVIEW; 642996); and Region Stockholm (ALF project). K.T. was supported by the Swedish Foundation for Strategic Research FFL18-0104 and D.L. by the China Scholarship Council. 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.

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

The PLR preprocessing workflow [implemented in the TimeStudio framework (Nyström et al., 2016)] and the statistical analysis scripts [implemented in R] are publicly available in OSF ([https://osf.io/kvepw/?view\\_only=ad68100a89034a0e8962cc1beaf34c2c](https://osf.io/kvepw/?view_only=ad68100a89034a0e8962cc1beaf34c2c)).

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

## Correspondence

Ana Maria Portugal, K6 Kvinnors och barns hälsa, K6 Neuropsykiatri Falck-Ytter, 171 77 Stockholm, Sweden; Email: ana.maria.portugal@ki.se

## Key points

- Pupil measures are closely linked to specific brain systems and implicated in autism and schizophrenia.
- For the first time, we employed a genetic-informed design to report on the relative contributions of genetic and environmental factors on pupil measures in infants, and on the associations of these measures with common genetic variants of autism and schizophrenia.
- We showed that pupil baseline and pupillary light reflex constriction amplitude were highly heritable at 5 months, but there was substantial independence between the genetic factors influencing each.
- Pupil baseline at 5 months was associated with common genetic variants of schizophrenia.

## References

- Aston-Jones, G., & Cohen, J.D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28, 403–450.
- Consortium, T. S. W. G. of the P. G., Ripke, S., Walters, J.T., & O'Donovan, M.C. (2020). Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *medRxiv*, 2020.09.12.20192922. <https://www.medrxiv.org/content/10.1101/2020.09.12.20192922v1>
- Daluwatte, C., Miles, J.H., Christ, S.E., Beversdorf, D.Q., Takahashi, T.N., & Yao, G. (2013). Atypical pupillary light reflex and heart rate variability in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43, 1910–1925.
- Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., ... & Neale, B.M. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*, 51, 63–75.
- DiCriscio, A.S., & Troiani, V. (2017). Pupil adaptation corresponds to quantitative measures of autism traits in children. *Scientific Reports*, 7, 6476.
- Ebitz, R.B., & Moore, T. (2017). Selective modulation of the pupil light reflex by microstimulation of prefrontal cortex. *Journal of Neuroscience*, 37, 5008–5018.
- Falck-Ytter, T., Hamrefors, L.I., Sanchez, M.S., Portugal, A.M., Taylor, M., Li, D., ... & Ronald, A. (2021). 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*, 24, 217–227.
- Fan, X., Miles, J.H., Takahashi, N., & Yao, G. (2009). Abnormal transient pupillary light reflex in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 1499–1508.
- Fattal, J., Brascamp, J., Slate, R., Lehet, M., Achtyes, E., & Thakkar, K. (2020). Abnormal pupil light reflex relates to negative symptom severity in schizophrenia. *Biological Psychiatry*, 87, S138.
- Fish, L.A., Nyström, P., Gliga, T., Gui, A., Begum Ali, J., Mason, L., ... & Jones, E.J.H. (2021). Development of the pupillary light reflex from 9 to 24 months: Association with common autism spectrum disorder (ASD) genetic liability and 3-year ASD diagnosis. *Journal of Child Psychology and Psychiatry*, 62, 1308–1319.
- Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C.-A., & Smoller, J.W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications*, 10, 1776.
- Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M., & Cohen, J.D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, 10, 252–269.
- Grove, J., Ripke, S., Als, T.D., Mattheisen, M., Walters, R.K., Won, H., ... & Børglum, A.D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, 51, 431–444.
- Hakarem, G., Lidsky, A., & Sutton, S. (1966). Pupillary reactions to single light pulses in psychiatric patients and normals. *Journal of Nervous and Mental Disease*, 153, 286–291.
- Hall, C.A., & Chilcott, R.P. (2018). Eyeing up the future of the pupillary light reflex in neurodiagnostics. *Diagnostics*, 8, 19.
- Howard, D.M., Adams, M.J., Clarke, T.-K., Hafferty, J.D., Gibson, J., Shirali, M., ... & McIntosh, A.M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, 22, 343–352.
- Jones, E., Mason, L., Begum Ali, J., van den Boomen, C., Braukmann, R., Cauvet, E., ... & Johnson, M.H. (2019). Eurosibs: Towards robust measurement of infant neurocognitive predictors of autism across Europe. *Infant Behavior and Development*, 57, 101316.
- Kercher, C., Azinfar, L., Dinalankara, D.M.R., Takahashi, T.N., Miles, J.H., & Yao, G. (2020). A longitudinal study of pupillary light reflex in 6- to 24-month children. *Scientific Reports*, 10, 1205.
- Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, 7, 18–27.
- Maggioni, E., Squarcina, L., Dusi, N., Diwadkar, V.A., & Brambilla, P. (2020). Twin MRI studies on genetic and environmental determinants of brain morphology and function in the early lifespan. *Neuroscience & Biobehavioral Reviews*, 109, 139–149.
- Mäki-Marttunen, V., Andreassen, O.A., & Espeseth, T. (2020). The role of norepinephrine in the pathophysiology of schizophrenia. *Neuroscience & Biobehavioral Reviews*, 118, 298–314.
- Maletic, V., Eramo, A., Gwin, K., Offord, S.J., & Duffy, R.A. (2017). The role of norepinephrine and its  $\alpha$ -adrenergic receptors in the pathophysiology and treatment of major depressive disorder and schizophrenia: A systematic review. *Frontiers in Psychiatry*, 8, 42.

- Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2015). The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 28–35.
- Murphy, P.R., Robertson, I.H., Balsters, J.H., & O'Connell, R.G. (2011). Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology*, *48*, 1532–1543.
- Neale, M.C., Hunter, M.D., Pritikin, J.N., Zahery, M., Brick, T.R., Kirkpatrick, R.M., ... & Boker, S.M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*, 535–549.
- Nyström, P., Falck-Ytter, T., & Gredebäck, G. (2016). The TimeStudio Project: An open source scientific workflow system for the behavioral and brain sciences. *Behavior Research Methods*, *48*, 542–552.
- Nyström, P., Gliga, T., Nilsson Jobs, E., Gredebäck, G., Charman, T., Johnson, M.H., ... & Falck-Ytter, T. (2018). Enhanced pupillary light reflex in infancy is associated with autism diagnosis in toddlerhood. *Nature Communications*, *9*, 1678.
- Nyström, P., Gredebäck, G., Bölte, S., & Falck-Ytter, T. (2015). Hypersensitive pupillary light reflex in infants at risk for autism. *Molecular Autism*, *6*, 10.
- Stahl, E.A., Breen, G., Forstner, A.J., McQuillin, A., Ripke, S., Trubetskoy, V., ... & Sklar, P. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, *51*, 793–803.
- Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I.I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*, 623–628.
- Wang, Y., Zekveld, A.A., Naylor, G., Ohlenforst, B., Jansma, E.P., Lorens, A., ... & Kramer, S.E. (2016). Parasympathetic nervous system dysfunction, as identified by pupil light reflex, and its possible connection to hearing impairment. *PLoS One*, *11*, e0153566.
- Zetterqvist, J., & Sjölander, A. (2015). Doubly robust estimation with the R Package drgee. *Epidemiologic Methods*, *4*, 69–86.

Accepted for publication: 18 November 2021