Heritability of Psychological Traits and Developmental Milestones in Infancy
A Systematic Review and Meta-analysis

Chloe Austerberry, MRes; Maria Mateen, DClinPsy; Pasco Fearon, PhD; Angelica Ronald, PhD

Abstract

IMPORTANCE  Although infancy is the most rapid period of postnatal growth and development, factors associated with variation in infant traits are not well understood.

OBJECTIVE  To synthesize the large twin study literature partitioning phenotypic variance in psychological traits and developmental milestones in infancy into estimates of heritability and shared and nonshared environment.

DATA SOURCES  PubMed, PsycINFO, and references of included publications were searched up to February 11, 2021.

STUDY SELECTION  Peer-reviewed publications using the classical twin design to study psychological traits and developmental milestones from birth to 2 years old were included.

DATA EXTRACTION AND SYNTHESIS  Data were extracted in line with Preferred Reporting Items for Systematic Reviews and Meta-Analyses and categorized using the International Classification of Functioning, Disability and Health: Children and Youth Version. Data were pooled in 3-level random effects models, incorporating within-cohort variance in outcome measurement and between-cohort variance. Data were analyzed from March 2021 through September 2021.

MAIN OUTCOMES AND MEASURES  The primary outcomes were monozygotic and dizygotic twin correlations. These were used to calculate genetic and shared and nonshared environment estimates.

RESULTS  Among 139 publications that were systematically retrieved, data were available on 79,044 twin pairs (31,053 monozygotic and 47,991 dizygotic pairs), 52 independent samples, and 21 countries. Meta-analyses were conducted on psychological traits and developmental milestones from 106 publications organized into 10 categories of functioning, disability, and health. Moderate to high genetic estimates for 8 categories were found, the highest of which was psychomotor functions (pooled $h^2$, 0.59; 95% CI, 0.25-0.79; $P < .001$). Several categories of traits had substantial shared environment estimates, the highest being mental functions of language (pooled $c^2$, 0.59; 95% CI, 0.24-0.86; $P = .001$). All examined categories of traits had moderate or high nonshared environment estimates, the highest of which were emotional functions (pooled $e^2$, 0.42; 95% CI, 0.33-0.50; $P < .001$) and family relationships (pooled $e^2$, 0.42; 95% CI, 0.30-0.55; $P < .001$).

CONCLUSIONS AND RELEVANCE  These findings may be an important source of information to guide future gene discovery research, public perspectives on nature and nurture, and clinical insights into the degree to which family history and environments may estimate major domains of infant functioning, disability, and health in psychological traits and developmental milestones.

Key Points

Question  What are the overall genetic and shared and nonshared environment estimates for psychological traits and developmental milestones in infancy?

Findings  In this systematic review and meta-analysis of 139 infant twin studies involving almost 80,000 twins globally, moderate to high genetic and shared and nonshared environment estimates were found across a range of important psychological traits and developmental milestones in infancy.

Meaning  These results offer insight into the degree to which genes and environments estimate outcomes in key domains of infant functioning and suggest highly heritable traits that may be particularly suitable candidates for gene discovery.


+ Supplemental content

Author affiliations and article information are listed at the end of this article.
Introduction

Infancy represents the most rapid period of postnatal growth and development, and research suggests that it is a critical or sensitive period for a wide range of psychological and developmental milestones. Investment in early childhood is argued to be one of the most effective economic strategies through promoting long-term socioeconomic and health outcomes. Investment before age 2 years, in particular, appears to be associated with the greatest rate of return for investment. This is reflected in an increasing policy focus globally on the first thousand and one days from conception to age 2 years.

Evidence suggests that complex traits are substantially but not entirely, heritable. Consequently, to gain understanding of the development of traits in infancy, it is important to draw on literature examining genetic and environmental factors associated with infant trait variation. The quantitative genetic method most widely and comprehensively performed in infancy is the classical twin design, which has been used for more than a century to partition phenotypic variance into additive genetic variance (heritability) and variance in the shared and nonshared environment. Family studies comparing biologically related siblings or parent-offspring pairs are typically unable to separate genetics from shared environment. In contrast, the classical twin design can provide separate estimates of heritability (the proportion of trait variation explained by genetic differences) and shared and nonshared environment. Twin studies are more feasible than adoption studies (which compare degree of resemblance between adoptees and their birth parents with resemblance between adoptees and their adoptive parents) to conduct at scale during infancy because adoption often occurs later in childhood. This has resulted in a far smaller and less comprehensive body of evidence in infancy from adoption studies than twin studies. The molecular genetic literature on infant traits is also small; the first genome-wide association study of infant traits was only recently conducted, and most molecular genetic studies in infancy have used candidate gene association methods, which in general have failed to yield replicable findings.

A landmark meta-analysis synthesizing virtually all twin studies of complex traits (predominantly psychiatric, metabolic, and cognitive traits) found a heritability of 49% across the lifespan when all traits and age groups were combined. The analysis combined data from infants and older children, calculating pooled estimates for children aged 0 to 11 years. Infancy is a rapid and sensitive period of development that deserves special focus. To address this, we conducted the first, to our knowledge, meta-analysis of twin studies of psychological and developmental functioning, disability, and health in infancy (birth to age 2 years), calculating pooled estimates of heritability and shared and nonshared environment.

Methods

This study protocol was registered with PROSPERO (record number, CRD42019151532), and the systematic review and meta-analysis were performed in line with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline 2020 statement and Meta-analysis of Observational Studies in Epidemiology (MOOSE) reporting guideline proposal for reporting. Given that the review involved the synthesis of anonymized information available in the public domain, it was exempt, according to University College London Research Ethics Committee (UCL REC) regulations, from requirements for ethics review by the UCL REC and the need for informed consent.

Search Strategy

PubMed and PsycINFO databases were searched on November 30, 2018; February 5, 2020; and February 11, 2021, for twin studies (a genetically informed design described in the eMethods in the Supplement) of psychological traits and developmental milestones in infancy, using the search terms in eTable 1 in the Supplement. Search results were imported into EndNote software version 9 (Clarivate). C.A. reviewed duplicates identified by EndNote, deleting true duplicates, and screened titles and abstracts of identified records against inclusion and exclusion criteria (eTable 2 in the Supplement). Full texts were retrieved for nonexcluded records, and these, along with references of
included publications, were screened by C.A. and M.M. Uncertainty about whether publications met inclusion criteria was resolved with senior researchers P.F. and A.R.

Quality Assessment and Data Extraction
Included publications were rated by C.A. using an adaptation for twin studies of the Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields for Quantitative Studies10 (eMethods in the Supplement). Information was extracted from each included publication by C.A. and M.M. (eTable 3 in the Supplement). If publications reported overlapping data, the estimate with the larger sample size (or, if sample sizes were identical, the most recently published estimate) was retained for meta-analysis (eMethods in the Supplement).

Classification of Phenotypes
Phenotypes were classified by C.A. using the World Health Organization International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY).11 Uncertainty about which ICF-CY category to use for a phenotype was resolved through discussion with M.M. and senior researchers P.F. and A.R. Phenotypes were excluded from the meta-analysis if they could not be categorized or were in categories containing data from fewer than 5 independent samples.

Statistical Analysis
In the R package metafor,12 we conducted two 3-level multilevel random-effects models (incorporating sampling variance, within-cohort variance in outcome measurements, and between-cohort variance) on twin correlations weighted by sample size from 10 categories of the ICF-CY containing data from 5 or more twin cohorts (eMethods in the Supplement). Zygosity was included as a moderator, with the dizygotic (DZ) group coded as the reference category in the first model to obtain a pooled monozygotic (MZ) twin correlation (MZ $r$) and standard error. The second model was identical but reparameterized with the MZ group as the reference category, producing a pooled DZ twin correlation (DZ $r$) and standard error. To allow for differences in variability in MZ and DZ subsets, models had a random error structure creating separate study-level and outcome error terms for MZ and DZ twins. As detailed in the eMethods in the Supplement, using pooled correlations and variances from the multilevel model, we calculated estimates for how much of the variation in each ICF-CY category was explained by additive genetic factors (A), the shared environment (C), and the nonshared environment (E, known collectively as ACE estimates) from standard univariate twin models estimated in the meta-analytic context using the R package metaSEM.13 We used 95% CIs around the pooled estimates in the twin study meta-analysis. Forest plots for analyses were produced using the R package metafor version 2.4-0 for R statistical software version 4.0.2 (R Project for Statistical Computing).12 We calculated $I^2$ for each of 3 levels in multilevel models. According to Cochrane guidelines, $I^2 \leq 40\%$ suggests low heterogeneity, while $I^2 = 30\%-60\%$ suggests moderate heterogeneity and $I^2 \geq 50\%$ indicates substantial or considerable heterogeneity.14 To reduce heterogeneity, analysis steps were repeated in 10 ICF-CY subcategories (with data from $\geq 5$ samples) and 3 ICF-CY categories (with separate data from parents and observers from $\geq 5$ samples) by parent and observer subgroup (for 6 meta-analyses in total) given that differences in rater have been found to be associated with differences in heritability estimates.15,16

We ran Egger tests of publication bias using the standard error as the estimator, and created funnel plots, plotting effect sizes against standard errors.17 In line with Cochrane recommendations, publication bias tests were run only on estimates in trait categories containing at least 10 estimates.14 Egger tests of publication bias were 2-sided and were considered significant at $P < .05$. Data analysis was conducted March through September 2021.
Results

We identified 5047 publications (4675 publications in databases and 372 publications in references). After duplicate removal and screening, 139 publications were included, containing data on 79,044 twin pairs (31,053 MZ and 47,991 DZ twins), 52 twin cohorts, 21 countries, and 6 continents between 1972 and 2020. The sample included 66,407 twin pairs from Western, educated, industrialized, rich, and democratic countries (in Europe, North America, and Oceania, 84.01%) and 12,637 twin pairs from Africa, Asia, and South America (15.99%) (eResults, eFigures 2 and 3, and eTable 4 in the Supplement). We extracted 2279 estimates (twin correlations or ACE estimates, including 1097 estimates from MZ twins and 1182 estimates from DZ twins) on 377 phenotypes, organized into 17 categories and 28 subcategories of the ICF-CY. Data from 33 publications included in the systematic review were excluded from the meta-analysis. Detailed information on search results, phenotype categorization, and excluded data is provided in eResults, eTables 1 and 4 through 6, and eFigures 1 through 3 in the Supplement.

<table>
<thead>
<tr>
<th>ICF-CY category</th>
<th>Example phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>b134 Sleep functions</td>
<td>Nocturnal sleep duration</td>
</tr>
<tr>
<td></td>
<td>Night awakenings</td>
</tr>
<tr>
<td></td>
<td>Sleep problems</td>
</tr>
<tr>
<td>b140 Attention functions</td>
<td>Attention problems</td>
</tr>
<tr>
<td></td>
<td>Task orientation</td>
</tr>
<tr>
<td></td>
<td>Spectral amplitude during visual attention</td>
</tr>
<tr>
<td>b147 Psychomotor functions</td>
<td>Activity level</td>
</tr>
<tr>
<td></td>
<td>Fine motor</td>
</tr>
<tr>
<td></td>
<td>Sitting without support</td>
</tr>
<tr>
<td>b152 Emotional functions</td>
<td>Resistance to soothing</td>
</tr>
<tr>
<td></td>
<td>Fearfulness</td>
</tr>
<tr>
<td></td>
<td>Positive affect</td>
</tr>
<tr>
<td>b163 Basic cognitive functions</td>
<td>General cognitive ability</td>
</tr>
<tr>
<td></td>
<td>Nonverbal cognitive development</td>
</tr>
<tr>
<td></td>
<td>Primary cognition</td>
</tr>
<tr>
<td>b167 Mental functions of language</td>
<td>Reception of language</td>
</tr>
<tr>
<td></td>
<td>Expressive vocabulary</td>
</tr>
<tr>
<td></td>
<td>Late language acquisition</td>
</tr>
<tr>
<td>b560 Growth maintenance functions</td>
<td>BMI</td>
</tr>
<tr>
<td></td>
<td>Head circumference</td>
</tr>
<tr>
<td></td>
<td>Weight gain</td>
</tr>
<tr>
<td>d710 Basic interpersonal interactions</td>
<td>Disregard for others</td>
</tr>
<tr>
<td></td>
<td>Reciprocal social behavior</td>
</tr>
<tr>
<td></td>
<td>Shyness</td>
</tr>
<tr>
<td>d720 Complex interpersonal interactions</td>
<td>Disruptive behavior</td>
</tr>
<tr>
<td></td>
<td>Peer aggression</td>
</tr>
<tr>
<td></td>
<td>Disregard for rules</td>
</tr>
<tr>
<td>d720 Family relationships</td>
<td>Attachment security</td>
</tr>
<tr>
<td></td>
<td>Dependence</td>
</tr>
<tr>
<td></td>
<td>Separation distress</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); ICF-CY, International Classification of Functioning, Disability and Health for Children and Youth.
**Meta-analysis Results**

**Analysis of Phenotypes by Category**

Among 10 categories of infant psychological and developmental functioning, disability and health displayed in Table 1 and defined in the ICF-CY,11 there were enough data from independent samples (≥5 samples) for meta-analysis. Results are reported in Table 2 and the Figure. Forest plots for these meta-analyses are reported in eFigures 4 through 13 in the Supplement. More twin samples used in these meta-analyses contained parent-reported data (cohort k = 22) than observer-rated data (cohort k = 12) (eResults and eTable 7 in the Supplement).

**Heritability**  
The highest heritability estimate was for psychomotor functions (pooled $h^2$, 0.59; 95% CI, 0.25-0.79; $P < .001$), followed by attention functions (pooled $h^2$, 0.48; 95% CI, 0.17-0.71; $P = .002$), complex interpersonal interactions (pooled $h^2$, 0.44; 95% CI, 0.15-0.75; $P = .003$), family relationships (pooled $h^2$, 0.41; 95% CI, 0.06-0.71; $P = .02$), and emotional functions (pooled $h^2$, 0.40; 95% CI, 0.16-0.64; $P = .001$). Remaining categories had lower estimates with 95% CIs above 0 (pooled $h^2$ range, 0.24-0.38), apart from mental functions of language and sleep functions, which had CIs overlapping 0 (pooled $h^2$, 0.24 and 0.35, respectively) (Table 2).

**Shared Environment**  
Mental functions of language (pooled $c^2$, 0.59; 95% CI, 0.24-0.86; $P = .001$), growth maintenance (pooled $c^2$, 0.46; 95% CI, 0.37-0.54; $P < .001$), basic cognitive functions (pooled $c^2$, 0.45; 95% CI, 0.21-0.69; $P < .001$), and sleep functions (pooled $c^2$, 0.45; 95% CI, 0.16-0.74; $P = .002$) had high shared environment estimates. Complex interpersonal interactions had a lower estimate (pooled $c^2$, 0.27; 95% CI, 0.04-0.51; $P = .02$), and estimates for psychomotor, attention, and emotional functions; family relationships; and basic interpersonal interactions had CIs overlapping with 0 (pooled $c^2$ range, 0.07-0.21) (Table 2).

**Nonshared Environment**  
Categories with the highest nonshared environment estimates were emotional functions (pooled $e^2$, 0.42; 95% CI, 0.33-0.50; $P < .001$), family relationships (pooled $e^2$, 0.42; 95% CI, 0.30-0.55; $P < .001$), basic interpersonal interactions (pooled $e^2$, 0.41; 95% CI, 0.30-0.52; $P < .001$), and attention functions (pooled $e^2$, 0.40; 95% CI, 0.29-0.51; $P < .001$). Remaining categories had lower estimates with CIs above 0 (pooled $e^2$ range, 0.18-0.33) (Table 2).

**Heterogeneity**  
Sampling variance contributed little to the total variance of each phenotypic category (level 1 $I^2$ range, 0.19%-12.44%) (Table 2). Within-cohort heterogeneity (ie, differences across measures within a domain and within a cohort) contributed substantially to total variance in mental functions of language, emotional functions, and growth functions (level 2 $I^2$ range, 58.59%-69.93%), and between-cohort heterogeneity contributed a lower amount (level 3 $I^2$ range, 24.73%-38.74%) to these outcomes. The remaining 7 categories each had substantial between-cohort heterogeneity (level 3 $I^2$ range, 56.56%-75.28%) and lower within-cohort heterogeneity (level 2 $I^2$ range, 23.59%-42.04%) (Table 2).

**Analysis of Phenotypes by Subcategory and Rater**

To reduce heterogeneity, we analyzed 10 subcategories of the ICF-CY (with data from ≥5 samples) and 3 phenotypic categories (with separate parent and observer data from ≥5 samples) by rater (for 6 subgroups: 3 with parent report and 3 with observer report). Full findings are reported in eResults and eTables 8 and 9 in the Supplement. Parent-rated phenotypes in the 3 examined categories (psychomotor and emotional functions and basic interpersonal interactions) had higher heritability and lower nonshared estimates than observer ratings and comparable shared environment estimates.
Table 2. Multilevel Random Effects Models of Phenotypic Categories

<table>
<thead>
<tr>
<th>ICF-CY category</th>
<th>Cohort ( k )</th>
<th>Estimates ( k )</th>
<th>MZ twin pairs, No.</th>
<th>DZ twin pairs, No.</th>
<th>Pooled MZ r (95% CI)</th>
<th>Pool DZ r (95% CI)</th>
<th>( h^2 ) (95% CI)</th>
<th>( e^2 ) (95% CI)</th>
<th>( f^2 ) (95% CI)</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep functions</td>
<td>7</td>
<td>49</td>
<td>1923</td>
<td>4044</td>
<td>0.80 (0.67-0.93)</td>
<td>0.63 (0.49-0.76)</td>
<td>0.35 (0.07-0.73)</td>
<td>.06</td>
<td>0.45 (0.16-0.74)</td>
<td>.002</td>
<td>0.20 (0.07-0.33)</td>
<td>.003</td>
</tr>
<tr>
<td>Attention functions</td>
<td>10</td>
<td>175</td>
<td>3011</td>
<td>6137</td>
<td>0.60 (0.49-0.71)</td>
<td>0.36 (0.25-0.47)</td>
<td>0.48 (0.17-0.71)</td>
<td>.002</td>
<td>0.12 (0.07-0.37)</td>
<td>.33</td>
<td>0.40 (0.29-0.51)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Psychomotor functions</td>
<td>13</td>
<td>151</td>
<td>3109</td>
<td>6105</td>
<td>0.67 (0.55-0.79)</td>
<td>0.37 (0.25-0.49)</td>
<td>0.59 (0.25-0.79)</td>
<td>.001</td>
<td>0.07 (0.03-0.15)</td>
<td>.60</td>
<td>0.33 (0.22-0.45)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Emotional functions</td>
<td>14</td>
<td>216</td>
<td>1756</td>
<td>3633</td>
<td>0.58 (0.50-0.66)</td>
<td>0.38 (0.30-0.46)</td>
<td>0.40 (0.16-0.64)</td>
<td>.001</td>
<td>0.18 (0.08-0.28)</td>
<td>.06</td>
<td>0.42 (0.34-0.50)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Basic cognitive functions</td>
<td>5</td>
<td>47</td>
<td>2636</td>
<td>5371</td>
<td>0.79 (0.68-0.89)</td>
<td>0.62 (0.51-0.73)</td>
<td>0.34 (0.04-0.64)</td>
<td>.03</td>
<td>0.45 (0.21-0.69)</td>
<td>&lt;.001</td>
<td>0.21 (0.11-0.32)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Language</td>
<td>5</td>
<td>96</td>
<td>2332</td>
<td>2853</td>
<td>0.82 (0.67-0.98)</td>
<td>0.71 (0.55-0.86)</td>
<td>0.24 (0.06-0.41)</td>
<td>.28</td>
<td>0.59 (0.24-0.86)</td>
<td>.001</td>
<td>0.18 (0.02-0.33)</td>
<td>.02</td>
</tr>
<tr>
<td>Growth</td>
<td>24</td>
<td>465</td>
<td>16653</td>
<td>21874</td>
<td>0.80 (0.76-0.83)</td>
<td>0.63 (0.59-0.67)</td>
<td>0.34 (0.23-0.45)</td>
<td>&lt;.001</td>
<td>0.46 (0.37-0.54)</td>
<td>&lt;.001</td>
<td>0.20 (0.17-0.24)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Basic interpersonal functions</td>
<td>18</td>
<td>356</td>
<td>4207</td>
<td>8037</td>
<td>0.59 (0.48-0.70)</td>
<td>0.40 (0.29-0.51)</td>
<td>0.38 (0.05-0.70)</td>
<td>.02</td>
<td>0.21 (0.08-0.48)</td>
<td>.10</td>
<td>0.41 (0.30-0.52)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Complex interpersonal functions</td>
<td>10</td>
<td>73</td>
<td>3244</td>
<td>5117</td>
<td>0.72 (0.61-0.82)</td>
<td>0.49 (0.39-0.60)</td>
<td>0.44 (0.15-0.75)</td>
<td>.003</td>
<td>0.27 (0.04-0.51)</td>
<td>.02</td>
<td>0.29 (0.18-0.39)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Family relationships</td>
<td>7</td>
<td>29</td>
<td>678</td>
<td>1546</td>
<td>0.58 (0.45-0.71)</td>
<td>0.37 (0.24-0.50)</td>
<td>0.41 (0.06-0.71)</td>
<td>.02</td>
<td>0.17 (0.04-0.45)</td>
<td>.24</td>
<td>0.42 (0.30-0.55)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Abbreviations: DZ, dizygotic; ICF-CY, International Classification of Functioning, Disability and Health, Children and Youth Version; MZ, monozygotic.

\( a \) Heterogeneity.

\( b \) Definitions for categories and subcategories can be found in cited ICF-CY manual.\(^1\)

\( c \) Number of independent twin cohorts.

\( d \) Number of estimates (twin correlations).

\( e \) MZ twin correlation.

\( f \) DZ twin correlation.

\( g \) Heritability.

\( h \) Shared environment.

\( i \) Nonshared environment.

\( j \) Sampling variance.

\( k \) Within-cohort variance in outcome measurement.

\( l \) Between-cohort variance.
Publication Bias
Possible publication bias was detected in the unexpected direction across all categories. Findings are in eResults, eTables 10 to 11, and eFigures 14 to 18 in the Supplement.

Quality Assessment
Quality assessment results are displayed in eFigure 19 in the Supplement. The mean (SD) score for the 106 publications included in the meta-analysis was 75.58% (13.83%).

Discussion
Drawing on a systematically-retrieved pooled sample of 79,044 twins, this systematic review and meta-analysis found evidence that most domains of functioning, disability, and health in psychological and developmental milestones were heritable in infancy and had moderate to high nonshared estimates. Contrary to evidence in older ages, shared environment estimates were high across several important domains of infant development.

Heritability
Consistent with evidence in older samples, all meta-analyzed categories had heritability estimates with 95% CIs above 0 in infancy, apart from sleep and language functions. Estimates were high (≥40%) for important areas of development (psychomotor, attention, and emotional functions; family relationships [attachment and dependency]; and complex interpersonal interactions [behavioral problems]), suggesting that these categories may be particularly suitable candidates for gene mapping.

High heritability in infancy of attention functions was consistent with the high heritability of attention-deficit/hyperactivity disorder (ADHD) and ADHD traits in older samples. In accordance

Figure. Ternary Plot of Pooled Heritability and Shared and Nonshared Environment Estimates by Phenotypic Category

Heritability
Psychomotor
Emotional
Sleep
Family relationships
Basic interpersonal
Behavioral problems
Basic cognitive
Attachment and dependency
Complex interpersonal
Language
Growth
Shared environment
Nonshared environment
Heritability
80%
60%
40%
20%
0%
80%
60%
40%
20%
0%

with the very high heritability of autism,\textsuperscript{158} a neurodevelopmental condition involving differences in social interaction, social cues in relationships, and regulating behaviors within interactions were among the most heritable subcategories. Absence of evidence that infant language was heritable was consistent with evidence that the heritability of cognition, including language, is low in early development, increasing with age.\textsuperscript{159,160}

The higher heritability of parent-rated than observer-rated phenotypes may be driven by contrast bias in parental reports of phenotypes among children who were DZ twins, exaggerating DZ differences, or by assimilation bias in parental reports of MZ twins, exaggerating MZ similarity.\textsuperscript{161} Correlated rater bias that inflated MZ and DZ twin similarity equally would lead to inflated shared environment estimates. Without raw data from individual studies, it was not possible to test this by examining variance-covariance structures, which can uncover evidence of contrast and assimilation bias. Overall, our results suggest that individual differences in growth, motor, cognitive, and emotional development may be associated with genetic factors as early as the first 2 years of life.

**Shared Environment**
Contrary to evidence in older age groups,\textsuperscript{7} shared environment estimates had CIs above 0 in several domains and were high for language, sleep, growth maintenance, and basic cognitive functions, reflecting a broader trend noted in the literature that shared environment estimates for language and cognition are higher in early development.\textsuperscript{159,160} This may have important implications for obesity prevention and efforts to promote intellectual outcomes, which are among the most robust estimators of health and longevity.\textsuperscript{162} Shared environment estimates had CIs overlapping with 0 for psychomotor, attention, and emotional functions; basic interpersonal interactions; and family relationships. This was consistent with pooled findings in older age groups\textsuperscript{7} and evidence that shared environments do not contribute as much to similarity between siblings as genetics and do not contribute as much to differences between siblings as nonshared environments.\textsuperscript{163,164}

**Nonshared Environment**
Nonshared environment estimates had 95\% CIs above 0 for all phenotypic categories and were high for emotional and attention functions, family relationships, and basic interpersonal interactions. Higher nonshared estimates for observer ratings than parent ratings were consistent with wider research\textsuperscript{165} and may reflect the importance of each twin's unique experience in the expression of phenotypes specifically when rated by observers. Alternatively, given that nonshared estimates also include measurement error, higher observer-rated estimates may reflect increased error in observational measurement.

**Limitations**
This study has several limitations. Given that research designs all have limitations and biases, establishing robust evidence ideally involves triangulation of methods. However, the classical twin design is currently the only quantitative genetic method that has produced data from enough independent samples to conduct adequately powered meta-analyses across a comprehensive range of infant traits. The generalizability of twin findings may be limited because some infant phenotypes (eg, language and birth weight) develop differently in twins compared with individuals from singleton births.\textsuperscript{166,167} However, given that our aim was to examine individual differences rather than investigate how and why groups differed, mean differences between twins and singletons may not indicate issues with generalizability.

Although the twin method can be used to examine genotype-environment correlation or interaction, we did not synthesize findings on these outcomes. In twin modeling, ignored interaction between genotype and shared environment is estimated as heritability and ignored interaction between genotype and nonshared environment will be estimated as nonshared environment, potentially contributing to biased estimates.\textsuperscript{168}
CIs for some estimates were wide. In meta-analysis, CIs depend on the precision of included studies, which are influenced by sample size and, in the case of twin modeling, the ratio of MZ to DZ pairs and relative contribution of each parameter. Furthermore, for any given sample size, there is more power to estimate \( e^2 \) than \( h^2 \) and \( c^2 \), which may explain the narrower intervals around \( e^2 \). CIs are also associated with the number of samples included in a meta-analysis; while adding studies may improve precision, it may also increase heterogeneity, decreasing precision. Crucially, in multilevel modeling, CIs are also dependent on the degree of between-study heterogeneity. High variability in estimates across studies is associated with wider CIs, and ignoring such heterogeneity tends to overestimate precision. The heterogeneity we observed was generally high, and so CIs were comparatively wide.

A downside of the comprehensive approach we took is that it may have increased between-study heterogeneity. We attempted to reduce this in the narrower subcategory and rater analyses. However, between-study heterogeneity was substantial in all categories and subcategories, suggesting that between-study differences likely created considerable noise in our estimates. Understanding and reducing heterogeneity will be important for future research to provide more precise twin estimates in infancy. Possible publication bias was also detected across multiple outcome domains. The impact of this on the estimates is difficult to rigorously assess.

Although individuals from Western, educated, industrialized, rich, and democratic societies represent 12% of the world's population, twins from these areas of the world constituted more than 80% of our sample. Infants in Africa, Asia, and South America combined represented approximately 16% of our sample, highlighting a need for twin research on infants in these continents.

There was an imbalance in the amount of research conducted on the synthesized categories; for example, far more was conducted on anthropometric phenotypes included in growth maintenance functions (which included data from 24 of 52 included samples) than other domains. Important areas in which research was lacking included nonsocial autistic traits and dysregulation, eating behavior, memory, higher-level cognitive functions, and brain structure.

**Conclusions**

To our knowledge, this systematic review and meta-analysis is the first study to synthesize the large and comprehensive infant twin literature on psychological traits and developmental milestones, offering insight into the possible earliest manifestations of phenotypic variance associated with genetic and environmental factors. This has the potential to improve public perceptions on nature and nurture by, for example, dispelling widely held beliefs that infants are shaped entirely by their environments or that family history entirely predetermines child health, beliefs that may place undue pressure on parents. For researchers, these results may offer a guide for future gene discovery research and efforts to uncover the causes of variation in infant traits. For clinicians, these findings may provide an indication of how family history and environmental conditions may estimate infant outcomes, including outcomes that may be early markers associated with subsequent healthy or pathological development.

**ARTICLE INFORMATION**

Accepted for Publication: July 5, 2022.


Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2022 Austerberry C et al. JAMA Network Open.

Corresponding Authors: Chloe Austerberry, MRes, Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom (c.austerberry@ucl.ac.uk); Angelica Ronald, PhD, Centre for


Downloaded From: https://jamanetwork.com/ on 09/17/2023
Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom (a.ronald@bbk.ac.uk).

Author Affiliations: Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom (Austerberry, Mateen); Centre for Family Research, University of Cambridge, Cambridge, United Kingdom (Fearon); Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom (Ronald).

Author Contributions: Ms Austerberry had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Drs Fearon and Ronald contributed equally to the manuscript.

Concept and design: Austerberry, Fearon, Ronald.

Acquisition, analysis, or interpretation of data: Austerberry, Mateen, Fearon.

Drafting of the manuscript: All authors.

Critical revision of the manuscript for important intellectual content: Austerberry, Mateen, Fearon, Ronald.

Statistical analysis: Austerberry, Mateen, Fearon.

Obtained funding: Austerberry.

Administrative, technical, or material support: Mateen.

Supervision: Fearon, Ronald.

Conflict of Interest Disclosures: Dr Fearon reported receiving an honorarium as deputy editor in chief of the Journal of Child Psychology and Psychiatry outside the submitted work. Dr Ronald reported receiving an honorarium as editor of the Journal of Child Psychology and Psychiatry and grants from the Genetics Society and Simons Foundation outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported by a studentship ES/P000592/1 from the UK Research and Innovation (UKRI) Economic and Social Research Council (ESRC) awarded to Ms Austerberry by the University College London, Bloomsbury, and East London Doctoral Training Partnership.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the organizations they are affiliated with or funded by.

REFERENCES


162. Deary IJ, Weiss A, Batty GD. Intelligence and personality as predictors of illness and death: how researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychol Sci Public Interest.* 2010;11(2):53-79. doi:10.1177/1529100610387081


SUPPLEMENT.

eMethods.
eResults.
eTable 1. Database and Reference List Searches
eTable 2. Inclusion and Exclusion Criteria
eTable 3. Information Extracted From Included Publications
eTable 4. Publications Identified in Systematic Literature Search, Presented Alphabetically by First Author
eTable 5. Twin Studies Identified in Systematic Literature Search, Presented Alphabetically by Study Name
Table 6. Phenotypes Identified in Systematic Literature Search Coded Using Classification System
eTable 7. Estimates and Cohorts in Phenotypic Categories by Rater
eTable 8. Multilevel Random Effects Models of Phenotypic Subcategories
eTable 9. Multilevel Random Effects Models of Parent and Observer Ratings of 3 Phenotypic Categories
eTable 10. Tests for Publication Bias on Twin Correlations by Phenotype Category
eTable 11. Tests for Publication Bias on Phenotype Categories by Variance Component in Phenotype Categories With ≥10 Estimates
eFigure 1. Prisma Flow Diagram
eFigure 2. Bar Chart of Number of Twin Pairs by Country
eFigure 3. Bar Chart of Number of Twin Pairs by Continent
eFigure 4. Sleep Functions Forest Plot
eFigure 5. Attention Functions Forest Plot
eFigure 6. Psychomotor Functions Forest Plot
eFigure 7. Emotional Functions Forest Plot
eFigure 8. Basic Cognitive Functions Forest Plot
eFigure 9. Mental Functions of Language Forest Plot
eFigure 10. Growth Maintenance Functions Forest Plot
eFigure 11. Basic Interpersonal Interactions Forest Plot
eFigure 12. Complex Interpersonal Interactions Forest Plot
eFigure 13. Family Relationships Forest Plot
eFigure 14. Funnel Plots of Association Between Monozygotic Twin Correlation and Standard Error in Phenotype Categories With ≥10 Estimates
eFigure 15. Funnel Plots of Association Between Dizygotic Twin Correlation and Standard Error in Phenotype Categories With ≥10 Estimates
eFigure 16. Funnel Plots of Association Between Heritability and Standard Error in Phenotype Categories With ≥10 Estimates
Figure 17. Funnel Plots of Association Between Shared Environment and Standard Error in Phenotype Categories With ≥10 Estimates
Figure 18. Funnel Plots of Association Between Nonshared Environment and Standard Error in Phenotype Categories With ≥10 Estimates
Figure 19. Bar Plot of Quality Assessment Scores
Figure 20. Sustaining Attention Forest Plot
Figure 21. Psychomotor Control Forest Plot
Figure 22. Organization of Psychomotor Functions Forest Plot
Figure 23. Regulation of Emotion Forest Plot
Figure 24. Range of Emotion Forest Plot
Figure 25. Expression of Language Forest Plot
Figure 26. Respect and Warmth in Interactions Forest Plot
Figure 27. Social Cues in Relationships Forest Plot
Figure 28. Regulating Behaviors Within Interactions Forest Plot
Figure 29. Interacting According to Social Rules Forest Plot
Figure 30. Psychomotor Functions (Observer Report) Forest Plot
Figure 31. Psychomotor Functions (Parent Report) Forest Plot
Figure 32. Emotional Functions (Observer Report) Forest Plot
Figure 33. Emotional Functions (Parent Report) Forest Plot
Figure 34. Basic Interpersonal Interactions (Observer Report) Forest Plot
Figure 35. Basic Interpersonal Interactions (Parent Report) Forest Plot
References.