**ENVIRONMENTAL-RISK (E-RISK) LONGITUDINAL TWIN STUDY**
**CONCEPT PAPER FORM**

Proposing Author: Jasmin Wertz
Author's affiliation, phone, and e-mail address: Duke University, +1(919) 613-0352
jasmin.wertz@duke.edu

Sponsoring Investigator (if the proposing author is a student, a post-doc or a colleague): Terrie Moffitt, Avshalom Caspi

Proposed co-authors: Terrie Moffitt, Avshalom Caspi PLUS

- E-Risk team: Louise Arseneault + interested others
- Dunedin Parenting Study team: Richie Poulton, Sandhya Ramrakha, Bob Hancox, Judith Sligo, Jay Belsky
- Duke Genomics Team and Dan Belsky
- WLS and HLS analysis team: Hexuan Liu, JC Barnes, Peter Tanksley
- Members of other collaborating cohort teams (i.e., PIs and data analyst – tbd)

Provisional Paper Title: Multi-cohort analysis of genetic associations with parental investment across a child's life

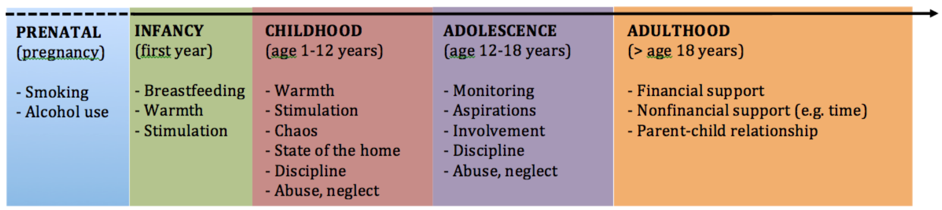
Date: 25 September 2019

Objective of the study and its significance:

**OBJECTIVE:** To test whether parents' polygenic scores for educational attainment are associated with parental investments across a child's life, from before a child is born (e.g. parental smoking, drinking during pregnancy), through infancy (e.g. breastfeeding), childhood (e.g. parental warmth; cognitive stimulation), adolescence (e.g. parental monitoring, involvement, aspirations), into adulthood (e.g. financial support).

**BACKGROUND:** The proposal builds on two previous studies (attached to this concept paper) in which we have shown that parents' and children's education-associated genetics predict parental investment in children's educational attainment. In both studies, education-associated genetics were operationalized using genome-wide polygenic scores, which are derived from genome-wide association studies (GWAS; Visscher et al., 2017) and aggregate millions of genetic variants across the genome into a score that indicates part of a person's genetic disposition to a particular trait or behavior (Dudbridge, 2013). In the first of these two studies (Wertz et al., in press, Developmental Psychology), we tested genetic associations with parenting in members of the Dunedin Study, a New Zealand-based population-representative birth cohort. We found that parents' education- associated genetics predicted warm, sensitive and stimulating parenting, both in personal interactions with their 3-year old children (as captured on video) and through the home environments they created for their families (as observed by home visitors). In the second study (Wertz et al., in press, Child Development), we extended this work to incorporate children's genetics in addition to mothers', to test genetic associations with parenting in 860 families participating in the Environmental Risk (E-Risk) Longitudinal Twin Study, a UK-based population-representative cohort. We found that both mothers' and children's education-associated genetics predicted parenting provided to children when they were between 5-12 years old. We also found that mothers' education polygenic scores predicted children's educational attainment when children were 18 years old, net of children's own education polygenic scores. This effect was mediated via the parenting mothers provided, particularly the cognitive stimulation of their children. <u>Here we propose to extend our previous work by adopting a lifespan view of parenting</u>**.** Our previous studies examined parenting during childhood, when children between age 3-12 years old. However, individual differences in parental investment emerge from before a child is born (e.g., in parental behaviors during pregnancy), and continue after childhood, through adolescence (e.g., in the form of parental monitoring) and into adulthood (e.g., in the form of financial transfers to children) (Figure 1). We propose to test whether parents' education-associated genetics are associated with these changing forms of parental investment across a child's life. To test this hypothesis, and to ensure replication across different study samples, we propose to include multiple cohorts (see attached Table of proposed cohorts – more cohorts or developmental phases within a cohort might be added depending on data availability). We selected datasets based on their availability of genetic information for at least one parent. If genetic data are additionally available for a second parent and/or children, the analyses would also test associations between these scores and parental investment too (see analysis section for a more detailed description of planned analyses). We plan to use the polygenic score for educational attainment, based on the most recent GWAS (Lee et al., 2018).

**Figure 1.** The figure shows the developmental periods in the child's life that the study proposes to focus on. For each developmental period, we listed examples of parental investments that we propose to examine (depending on the availability of measures for these constructs in the participating cohorts).




**SIGNIFICANCE:** Studying correlations between genetics and parenting is significant for at least five reasons. First, studying gene-parenting correlations informs the interpretation of findings from studies that test effects of parenting on children's attainment. If the same genes that influence educational attainment also predict the kind of parenting that is linked with educational success, genetic influences may create the false impression of a causal relationship between parenting and children's attainment (Knafo & Jaffee, 2013). Second, studying gene-parenting correlations informs the interpretation of findings from genome-wide association studies of educational attainment. If education-associated genetics identified in GWAS partly reflect the quality of parenting that children receive, environmental influences may create the false impression of a causal relationship between genes and educational attainment (Young et al., 2018). Third, studying gene-parenting correlations can help us understand how genetics contribute to the continuity of behaviors across generations. If children's genes are correlated with the parenting they receive, it means that genetics may contribute to intergenerational continuity not only directly (via genetic inheritance) but also indirectly via nurture (i.e., by affecting the caregiving environment that shapes children's outcomes) (Kong et al., 2018). Fourth, studying gene-parenting correlations can help reveal how genetic variation is translated into behavior. If genetic influences on children's educational attainment partly manifest through shaping the parenting that children experience, it implies that environments are part of the pathway from genotype to phenotype (Belsky & Harden, 2019). Combining genetic data with measures of individuals' social environments is key to tracing such environmentally-mediated genetic effects. Fifth, studying gene-parenting correlations contributes to a more socially-informed view of genetics. If children's outcomes are influenced by the genetically-influenced behaviors of their parents, it broadens the scope of the study of genetics, from an individual's genes and their effects on an individual's phenotype, to the genome of people within an individual's social context (Domingue & Belsky, 2017).

Statistical analyses:

*Note:*

- *All of the following proposed analyses will use the most recent polygenic score for educational attainment (Lee et al., 2018).*
- *The exact variables to be included from each cohort is tbd depending on the constructs that have been measured and can be harmonized across cohorts.*

**Main Aim 1: Test associations between parental polygenic scores for educational attainment and parental investment.**

In each cohort, the main aim (Aim 1) will be to test associations between parental polygenic scores for educational attainment (Lee et al., 2018) and parental investment. The form of parental investment will vary across a child's life. From our review of available studies so far, we propose analyzing the constructs outlined in Table 1. Exact measures will differ across cohorts, depending on what data have been collected, although we will attempt to harmonize measures as much as possible. The statistical analysis to be used here would be a multiple linear or Poisson regression (depending on how parental investments has been measured), to test whether parental polygenic scores predict parental investments, adjusting for child sex as a covariate.

Most datasets have polygenic scores for only one parent. However, *if polygenic scores are available for both mum and dad*, we would test associations for both parents. We would also test assortative mating between parents (by testing whether education polygenic scores are correlated between parents) and additive contributions of mothers' and fathers' genetics on parental investment (by testing whether including adding the second parents' polygenic score when predicting parental investment adds to the prediction).

**Table 1.** Parental investment constructs that we propose to analyze, listed within each of the developmental phases of a child's life (exact measures will depend on data availability in each cohort).

| DEVELOPMENTAL PHASE IN CHILD'S LIFE | PROPOSED MEASURES (will depend on what is available in each cohort) |
|---|---|
| Prenatal | Substance use (cigarette, alcohol, drugs) |
| Infancy (first year) | Breastfeeding<br>Warm, sensitive parenting<br>Cognitive stimulation |
| Childhood (age 1-12) | Warm, sensitive parenting<br>Cognitive stimulation<br>Chaos in the home<br>State of the home<br>Parental discipline<br>Abuse and neglect |
| Adolescence (12-18) | Warm, sensitive parenting<br>Cognitive stimulation<br>Parental aspirations for child<br>Parental monitoring<br>Parental involvement<br>Parental monitoring<br>Abuse and neglect |
| Adulthood (>age 18) | Financial support (wealth transfers)<br>Nonfinancial support (time investments)<br>Quality of parent-child relationship |

**Optional aim 2: Only *if polygenic scores for educational attainment for children are available*, test associations between children's polygenic scores and parental investment.**

Not all cohorts will have genotyped children in addition to parents (for example, the Dunedin cohort has genotyped one parent, no children). However, if polygenic scores are available for children, it will be possible to test unique associations of parent and child education polygenic scores with parenting, by adding children's polygenic scores to a model predicting parental investments from parents' polygenic scores. This analysis tests for evocative effects of children's genetics on the parenting they receive. Research has shown that children influence the parenting they receive from their parents (for example, brighter children tend to receive more cognitive stimulation from their parents). We also observed in our previous research that children's polygenic scores predicted parental investments over and above parents' polygenic scores (Wertz et al., in press, Child Development). Here we propose to test the same hypothesis. The statistical analysis would proceed in three steps. First, using multiple linear or Poisson regression analyses as in Aim 1, test whether child polygenic scores predict parental investments, adjusting for child sex. Second, using simple correlation analyses, test whether parent and child polygenic scores are correlated (in biological parent-child dyads, they should be correlated at around r=.50 because parents pass on genes to their biological children). Third, using multiple linear or Poisson regression analyses, include both child and parent polygenic scores in the model to predict parental investments, adjusting for child sex.

**Optional aims 3 and 4: Only *if there is data on children's educational attainment (at any age, but the later, the better)* and *if parental investment has been assessed before children completed their education*, test whether parental investment predicts children's educational attainment (aim 3) and whether parental polygenic scores predict children's educational attainment (aim 4a) and whether effects of parental polygenic scores on children's educational attainment are mediated by parenting (aim 4b).**

The purpose of these analyses is to test the hypothesis that parental genetics do not only influence children's educational attainment via genetic transmission from parent to child, but also via the parental investments that parents make – an environmentally-mediated genetic effect (Kong et al., 2018; Bates et al., 2018; Belsky et al., 2018; Wertz et al., in press, Child Development). To test this, it is necessary to test whether parental education polygenic scores predict children's educational attainment, ideally adjusting for children's own education polygenic scores (if available), and if this effect (if evident) is mediated by parental investments. This analysis would only be applicable if the measure of parental investments in the study was collected before children finished their education (because only then is it reasonable to assume that parental investments had an effect on children's educational attainment). So, if parental investments have been measured in adulthood (say, in the form of financial transfers from parents to adult children), this test would not apply. The statistical analysis would proceed as follows: first, using multiple linear or Poisson regression analyses (depending on how children's educational attainment was measured), test whether parental investments predict children's educational attainment, adjusting for children's sex. Second, using multiple linear or Poisson regression analyses, test whether parental polygenic scores predict offspring's educational attainment, adjusting for children's sex. Third, if the effect of parental polygenic scores on child educational attainment is significant, add children's own polygenic score (if available) to the model, to test whether parents' polygenic scores predict attainment over and above children's own polygenic scores, adjusting for children's sex. Fourth, if the effect of parental polygenic scores is still significant, add measure(s) of parental investments to the model to test whether the effect of parental polygenic scores on child attainment is mediated by parental investments. To test mediation, conduct a formal test of mediation.

*If genotype data has been collected on both mother and father and at least one child*, it will be possible to construct polygenic scores based on non-transmitted versus transmitted alleles from parent to child (Kong et al., 2018). Specifically, by using genotype data from mother/father/child trios, it is possible to disentangle the alleles that have been passed on from parent to child, from the alleles that have not been passed on, exploiting the fact that parents only pass a random half of their genes on to their children. It is then possible to create a polygenic score based on only the alleles that have not been transmitted from parent to child. Using this score provides a strong test of the hypothesis that parents' genetics shape children's outcomes via something the parent does rather than by genetic transmission from parent to child (because these alleles have not been transmitted to the child). This logic is comparable to using an adoption design, where parents and children are not related, and therefore, any effect of adoptive parents' genes on their children's behavior is bound to be mediated by parental behavior.

Variables Needed at Which Ages (names and labels):

*Note: The decision about which variables will ultimately be used will be made based on a) harmonization across cohorts in this multi-cohort study; b) sufficient N for categories of variables that measure a rare outcome, e.g. neglect and abuse).*

Study: E-Risk

| PHASE | VARIABLE LABEL | VARIABLE DESCRIPTION (INFORMANT) |
|---|---|---|
| E-Risk cross-phase or no specific phase | FAMILYID | Family ID |
| | ATWINID | Twin ID |
| | SAMPSEX | Sex of child |
| | MOM_ZRPGSEA3 | Mom education polygenic score |
| | TWIN_ZRPGSEA3 | Child education polygenic score |
| | HARME512 | Child harm (parent, interviewer) |
| | PNSEVERITYE12 | Neglect (parent) |
| E-Risk Phase 1st contact | SMKGP | Smoking during pregnancy (mother) |
| | ABRSTFEDE | Twins breastfed (mother) |
| E-Risk Phase 5 | ACTVM5 | Activities with mother (Mother report) |
| | BP20EM5 | Maternal joy (interviewer) |
| | BP22EM5 | Maternal warmth (interviewer) |
| | BP24EM5 | Maternal affection (interviewer) |
| | WARME5 | Expressed emotion: warmth (mother) |
| | DISSE5 | Expressed emotion: dissatisfaction/negativity |

| | | |
|---|---|---|
| | TOTPHYE5 | Corporal punishment (mother) |
| E-Risk Phase 7 | HOMEM7<br>CHSTIM7<br>PAPSE7<br>PANGE7<br>CHAOSM7<br>NEGLCE7 | State of the home (interviewer)<br>Stimulation of the child (interviewer)<br>Positive parenting (interviewer)<br>Negative parenting (interviewer)<br>Chaos in the home (interviewer)<br>Child neglect (interviewer) |
| E-Risk Phase 10 | HOMEM10<br>CHSTIM7<br>PAPSE7<br>PANGE7<br>CHAOSM7<br>HAPPHM10<br>NEGLCE7<br>WARME10<br>DISSE10<br>SEPARATE ITEMS FOR HEALTH PARENTING | State of the home (interviewer)<br>Stimulation of the child (interviewer)<br>Positive parenting (interviewer)<br>Negative parenting (interviewer)<br>Chaos in the home (interviewer)<br>Happy home (interviewer)<br>Child neglect (interviewer)<br>Expressed emotion: warmth (mother)<br>Expressed emotion: dissatisfaction (mother)<br>Health parenting items (mother) |
| E-Risk Phase 12 | CHAOSEC12<br>HOMEM12, HOMEC12<br>CHSTIM12<br>CHAOSM7, CHAOSC7<br>HAPPHM12<br>NEGLCE7<br>ADULTEC12<br>MONE12, MONEM12, MONEP12<br>KNOWE12, KNOWEM12, KNOWEP12<br>CONTE12, CONTEM12, CONTEP12 | Home Chaos (child)<br>State of the home (interviewer)<br>Stimulation of the child (interviewer)<br>Chaos in the home (mother interviewer, child interviewer)<br>Happy home (interviewer)<br>Child neglect (interviewer)<br>Adult involvement (child)<br>Parental monitoring (child, mother, father)<br>Parental knowledge (child, mother, father)<br>Parental control (child, mother, father) |
| E-Risk Phase 18 | EDUCACHVE18<br>VCTZFAME18<br>VCTZNEGE18<br>VCTZMALE18 | Educational attainment (child)<br>Family victimization (child)<br>Neglect victimization (child)<br>Maltreatment victimization (child) |

Study: Dunedin (*Note: some of the analyses we propose will have already been presented in our Wertz et al., in press, Developmental Psychology paper. The difference is that we propose to add variables (i.e. discipline strategies) and that we plan to present the analyses as part of a larger picture of lifespan parenting. In the study, we'll clearly state which results have previously been reported (e.g. genetic associations with warm, sensitive, stimulating parenting) and which results will be new (e.g. proposed testing of genetic associations with discipline strategies)*.

| PHASE | VARIABLE LABEL | VARIABLE DESCRIPTION (INFORMANT) |
|---|---|---|
| Parenting Study Age 3 | snum<br>sex<br>parage<br>childage<br>pi1<br>pi6<br>pi10<br>pi17<br>vidratp3_sensitive_2016<br>vidratp3_intrusive_2016 | Study member ID<br>sex of parent<br>Age of parent<br>Age of child<br>Biological parent<br>Ordinal position of target child<br>Parental educational qualifications<br>Total household income<br>Video Rating parent sensitivity<br>Video Rating parent intrusiveness |

| | | | |
|---|---|---|---|
| | vidratp3_detach_2016<br>vidratp3_cogstim_2016<br>vidratp3_posreg_2016<br>vidratp3_negreg_2016<br>vidratc3_posmood_2016<br>vidratc3_negmood_2016<br>vidratc3_activitylvl_2016<br>vidratc3_persistence_2016<br>kidposv3_2016<br>kidnegv3_2016<br>parposv3_2016<br>Total HOME score<br>HOME – warm, sensitive parenting<br>HOME – cognitive stimulation<br>Discipline strategies (variables do not appear to be constructed yet) | Video Rating parent detachment<br>Video Rating parent cog stimulation<br>Video Rating parent positive regard<br>Video Rating parent negative regard<br>Video Rating child positive mood<br>Video Rating child negative mood<br>Video Rating child activity level<br>Video Rating child persistence<br>Video rating child positivity<br>Video rating child negativity<br>Video rating parent positivity<br>HOME variables as previously constructed for Wertz et al., in press Developmental Psychology<br>Discipline strategies questionnaire | |
| From main study | age_at_birth1<br>zrpgsEA3 | Parent age at first birth<br>Residualized standardized polygenic score for educational attainment | |

References cited:


Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., ... & Moffitt, T. E. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, *115*, E7275-E7284.

Belsky, D. W., & Harden, K. P. (2019). Phenotypic annotation: Using polygenic scores to translate discoveries from genome-wide association studies from the top down. *Current Directions in Psychological Science*, *28*, 82-90.

Domingue, B. W., & Belsky, D. W. (2017). The social genome: Current findings and implications for the study of human genetics. *PLoS Genetics*, *13*, e1006615. https://doi.org/10.1371/journal.pgen.1006615

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, *9*, e1003348. https://doi.org/10.1371/journal.pgen.1003348

Knafo, A., & Jaffee, S. R. (2013). Gene-environment correlation in developmental psychopathology. *Development and Psychopathology*, *25*, 1–6.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*, 424–428.

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: Biology, function, and translation. *American Journal of Human Genetics*, *101*, 5–22.

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., ..., & Caspi, A. (in press). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development*. (ATTACHED TO THIS CONCEPT PAPER)

Wertz, J., Belsky, J., Moffitt, T.E., Belsky, D.W., Harrington, H., Avinun, R., Houts, R., Poulton, R., Ramrakha, S., & Caspi, A. (in press). The genetics of nurture: A test of the hypothesis that parents' genetics predict observed caregiving behavior. *Developmental Psychology*. (ATTACHED TO THIS CONCEPT PAPER)

Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., … Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, *50*, 1304–1310.

**Table 2.** The table describes the cohorts that we have identified so far for inclusion in the proposed study. More cohorts (or data from additional developmental phases within a cohort) might be added depending on data availability.

| Cohort | Developmental phase of the child's life that is covered by the cohort | Study members (genotyped = bold) | Measures of parental investment |
|---|---|---|---|
| E-Risk cohort |  |  | Prenatal: Smoking<br>Infancy: Breastfeeding<br>Childhood:<br>• Warm, sensitive parenting<br>• Cognitive stimulation<br>• Household chaos<br>• State of the home<br>• Discipline strategies<br>• Abuse, neglect<br>• Health parenting<br>Adolescence:<br>• Parental monitoring<br>• Abuse, neglect |
| Millennium Cohort Study |  |  | Prenatal: Smoking, Drinking<br>Infancy: Breastfeeding<br>Childhood:<br>• Warm, sensitive parenting<br>• Cognitive stimulation<br>• Discipline strategies<br>• Health parenting<br>Adolescence:<br>• Parental monitoring<br>• Parental aspirations |
| Dunedin cohort |  |  | Childhood:<br>• Warm, sensitive parenting<br>• Cognitive stimulation<br>• Discipline strategies |

*Table 2 continued*

| | | | |
|---|---|---|---|
| Health and Retirement Study |  |  | Adulthood:<br>• Financial transfers<br>• Time investments |
| Wisconin Longitudinal Study |  |  | Adulthood:<br>• Financial transfers<br>• Time investments |

*Note:* Developmental periods are color-coded to match Figure 1: blue= prenatal period; green=infancy (<1 years of age); red=childhood (1-12 years of age); purple=adolescence (12-18 years of age); orange=adulthood (>18 years of age).
 = mother; = father; = child;  = mother or father (in some cohorts, only one parent was genotyped).