Data Supplement for Halldorsdottir et al.,Neurobiology of Self-Regulation: Longitudinal Influence of FKBP5 and Intimate Partner Violence on Emotional and Cognitive Development in Childhood. Am J Psychiatry (doi: 10.1176/appi.ajp.2019.18091018)

Supplemental Methods

Procedures

Families were visited in their homes at child ages 7, 15, 24, 36, 48 and 60 months and the primary caregiver (the mother in almost all instances) answered questions about demographics (e.g., child race, sex, maternal education, household size and income) and intimate partner violence. In addition to a number of other procedures, children were administered emotion induction tasks at 7, 15, and 24 months to elicit fear and frustration responses. At 36, 48 and 60 months, children were administered a battery of six tasks to assess EF. Children were also seen in school at prekindergarten, kindergarten, first, second, and fifth grade and administered measures of academic ability. Children's teachers at each grade in school completed questionnaires on child behavior.

Measures

Fear-Elicited Emotional Reactivity. Peak arousal was determined based on guidelines established in the experimental protocol. Behavioral reactivity during the stress paradigm was coded second by second from videotapes. A composite score for heightened reactivity to the stressor was created by summing the seconds of low, medium, and high reactivity. The proportion was then calculated by dividing this sum of heightened reactivity by the total duration of the task. Coders were trained to achieve 0.75 (Cohen's Kappa) reliability. Inter-rater reliability, assessed for at least 15% of cases, was 0.94, 0.89, and 0.90 at 7, 15, and 24 months, respectively.

Prolonged Stress-Induced Cortisol. Saliva was collected prior to and following the administration of a validated stress paradigm(1–3), including the mask presentation challenge(4). Saliva was collected prior to the administration of the stress paradigm and then again 20 and 40 minutes after the infant's peak emotional arousal to the stress paradigm. For the majority of the infants, peak arousal occurred following the administration of all the tasks; however, children who became highly aroused (i.e., more than 20 seconds of hard crying) during the course of the task administration were considered to have reached peak arousal.

Cotton or hydrocellulose absorbent material was used to collect unstimulated whole saliva and placed into 2 ml cryogenic storage vials using a needleless syringe (cotton) or by centrifugation (hydrocellulose) (5, 6). Samples were then placed on ice and stored frozen (–80 °C) until assay. The samples were assayed for salivary cortisol with the enzyme immunoassay U.S. Food and Drug Administration 510k cleared for use as an in vitro diagnostic measure of adrenal function (Salimetrics, State College, Pennsylvania). Finally, cortisol distributions were log transformed prior to analyses to correct for positive skew.

.

Executive Function. Each task was presented in an open spiral bound flipbook with pages that measured 8" x 14". For each task, trained research assistants established that the child knew colors and numbers and administered training trials and up to three practice trials if needed. If children failed to demonstrate an understanding of the goals of the task following the practice trials, the examiner discontinued that task. To capture the differing aspects of executive function, toddlers completed six tasks: Working Memory Span, Pick the Picture, Spatial Conflict Arrows, Something's the Same, Silly Sounds Stroop, and Animal Go No-Go. In brief, in the Working Memory Span, the child is presented with a picture of an outline of a house with an animal figure above which is a colored dot inside the house. The child is asked to name and hold in mind two pieces of information simultaneously and activate an animal name while overcoming interference occurring from naming the color of the dot presented earlier.

In Pick the Picture, the child is presented with a set of 2-, 3-, 4- or 6-pictures. For each picture set, the child is instructed to pick a picture from a series of picture so that each of the pictures from the set "gets a turn", i.e., that each picture is selected. In the Spatial Conflict Arrow, the child is presented with sequential pages on which an arrow at the top of the page is pointing to the left or right above one of two circles located at the lower left and lower right of the page, respectively. Task difficulty is manipulated by having the direction of the arrow either congruent or incongruent with the location of the response. During the Something's the Same, the child is instructed to identify similarities (e.g., shape, color, size) between pictures. In the Silly Sounds Stroop, the child is asked to produce an animal sound (e.g., sound of a cat) that does not fit with the presented picture (e.g., picture of a dog). During the Animal Go No-Go task, the child is instructed to click a buzzer when they see an animal with the exception of when that animal is a pig. Full details regarding the administration rules, psychometric properties, and scoring approach for each of these tasks have been presented elsewhere(7).

Full details regarding the administration rules, psychometric properties, and scoring approach for each of these tasks have been presented elsewhere(7).

Genotyping

In brief, 40 ng of DNA was combined in a volume of 5 µl with 2X Universal PCR Mix (Applied Biosystems) and 1/20 the volume of the Taqman SNP assay in a 384 well plate. A Pre-Read was performed and then PCR as follows: a 10 min hold at 95 °C, followed by 40 to 45 cycles of 15 s at 92 °C and then 1 min at 60 °C in a 7900HT PCR System. After amplification, a Post-Read was performed to analyze. Automatic and manual calls were made (8).

SNPs were quality controlled using procedures outlined previously (9); briefly, quality control required Hardy-Weinberg equilibrium testing $p<0.001$, missingness by marker <5%, missingness by sample <5%, affirmative relationship checking in PEDCHECK, and Mendelian inconsistency caused genotypes to be dropped at that locus.

Also genotyped was a panel of 48 SNPs that were chosen to include markers that provide information on both sample identification and relatedness to family members (highly polymorphic across the population) and ancestrally informative markers (polymorphic across human populations). This panel was validated for continental populations and also specifically for quantifying admixture in African American samples, as we have described previously (10).

Genotyping was conducted using a custom ligation detection reaction with a probe specifically tagged to identify each allele in a multiplex reaction (9, 11). We analyzed population ancestry using principal components analysis (PCA) calculated in EIGENSTRAT that computes PCA scores from SNP data; specifically in this application, 48 ancestrally informative SNPs validated for continental populations and quantifying admixture in African American samples (10). We analyzed all samples that passed quality control together with all 1000 Genomes samples with >85% of the 48 SNPs in the panel to provide clear reference populations for the three major continental groupings. The first three principal components were visualized graphically for our samples along with the 1000 Genomes samples, all color-coded by ancestry. African American samples displayed variation in the degree of admixture, as expected. As many of these samples did not clearly fall into either continental population according to Price et al. (12), it was necessary to use the quantitative information on ancestry to control for this potentially confounding variation. Therefore, we included the first 6 principal components to control for the effects of admixture in all analyses.

Statistical Analyses

Missing data. Table S1 displays descriptive statistics and number of participants for each variable in the analysis. Participants with missing data were generally similar to the analysis sample. Participants with missing IPV data were more likely to be African American ($r=0.14$, $p<0.0001$), and to be characterized by higher cumulative risk ($r=0.28$, $p<0.0001$). Missing data was imputed using the *missForest* package v1.4 (13). The imputed data was used in the longitudinal cluster analyses and mixed linear growth models.

Longitudinal Cluster Analyses. Consistent with the aforementioned analyses, the following covariates were used in the analysis: six multidimensional scaling components of the genetic relationship matrix to account for population structure, child sex, state of residency, and cumulative risk scores (plus acetaminophen intake in the prolonged cortisol analyses). The outcome variables were corrected for these covariates in a linear model and the residuals of this model were used for the longitudinal clustering. For each cluster analysis, the package *kml3d* v2.4.2 (14) in R v3.4.3 was used to assign individuals to homogeneous subgroups, jointly based on the outcomes of interest, while accounting for the longitudinal pattern over three time points. For a given number of subgroups (2–6), partitioning was optimized by maximizing the between-cluster variance. The Cohen's Kappa test, which was used to examine the correspondence between the clusters, is a conservative measure that takes into account the possibility of agreement between cluster occurring by chance.

Mixed linear growth models. Continuous predictor and moderator variables were grand mean centered. Child age at assessment was used as the within-person time variable. In all models, fixed effects included the indicators for CATT haplotype, IPV exposure, and time, along with the covariates and random intercepts and random linear slopes were estimated. As a sensitivity check, in all models we also examined the interaction of the *FKBP5* haplotype with the cumulative risk and chaos variables to confirm that the effect is specific to IPV exposure. The growth models were run in Mplus Version 7.2 and R using maximum likelihood estimation with robust standard errors.

Sensitivity Analyses. First, to exclude alternative explanations, we included intellectual ability, as measured by the MDI and the WPPSI full scale IQ, as a covariate in the IPV×CATT haplotype predicting reading and math skills. Secondly, to exclude maternal depressive symptoms as a confounder, we included the CES-D as a covariate in the model with the interaction between IPV and CATT haplotype predicting cluster group assignment. Third, to exclude ethnicity as a confounding variable of the findings, a three-way interaction of IPV and CATT haplotype with ethnicity/race on cluster group assignment was conducted using the ordinal regression. To determine if the effects were a by-product of maternal genotype, the interaction between maternal CATT haplotype and IPV exposure on childhood outcomes was explored. We also examined the interaction of the *FKBP5* haplotype with the cumulative risk and chaos variables on all outcomes to confirm that the effect is specific to IPV exposure. Lastly, we examined the effect of *FKBP5* haplotype×IPV on intellectual ability (not conceptualized within self-regulation) to explore whether this G×E effect was specific to self-regulation outcomes. See Figure S4 in supplement for a flow chart of the analyses.

Supplemental Results

In the early development analyses, 51.8% of the participants were within Cluster A and 48.2% in Cluster B. Cluster B was characterized by higher fear-elicited emotional reactivity across the time points than Cluster A (Figures S4A and S4B). For the school-aged outcomes, 65.2% of the participants were in Cluster C and 34.8% in Cluster D. Cluster C was characterized by less emotional and behavioral difficulties and high reading abilities across time than Cluster D (Figures S5A and S5B).

Mixed Linear Growth Models

Prolonged stress-induced cortisol. The interaction between IPV exposure and CATT haplotype did not significantly predict trajectories in prolonged stress-induced cortisol from 7 months, 15 months and 24 months (β=0.050, SE=0.028, p=0.074). However, at 15 and 24 months, the interaction of the CATT haplotype with IPV predicted prolonged stress-induced cortisol reactivity (β=0.10, SE=0.05, p=0.035 and β=0.09, SE=0.04, p=0.038 respectively). At both 15 and 24 months, carriers of two copies of the CATT haplotype exposed to high IPV displayed higher levels of prolonged stress-induced cortisol reactivity compared to those with one or no CATT haplotype copies (ES=0.13 at 15 months and ES=0.20 at 24 months).

Fear-Elicited Emotional Reactivity to the Mask Presentation. The IPV and CATT haplotype interaction did not predict the slope of change in fear-elicited emotional reactivity from 7, 15, and 24 months (β=0.046, SE=0.029, p=0.122). At child age 24 months, however, the interaction of CATT haplotype with caregiver report of IPV was associated with increased fear-elicited emotional reactivity to the mask presentation (β=0.06, SE=0.03, p=0.026), with the carriers of two copies of the CATT haplotype and exposure to IPV being the most reactive (ES=0.26).

Executive Function. The interaction between CATT haplotype and IPV exposure did not significantly predict trajectories in EF across 36, 48 and 60 months (β=–0.071, SE=0.062, p=0.248). Further inspection of the interaction revealed that the CATT haplotype interacted with

.

IPV to predict EF at 60 months (β=–0.079, SE=0.036, p=0.029), with carriers of two copies of the CATT haplotype exposed to high IPV exhibiting the lowest EF (ES=0.29).

Emotional and Behavioral Problems. A mixed linear model predicting teacher reported emotional and behavioral problems at grades 1, 2, and 5 indicated that the interaction of *FKBP5* genotype with exposure to IPV was associated with the intercept (β=0.048, SE=0.021, p=0.020) but not with linear or quadratic growth. This effect is present at each time point and not affected by the placement of the intercept. This analysis indicates that carriers of two copies of the CATT haplotype exposed higher levels of IPV exhibited higher levels of teacher-reported emotional and behavior problems in the early elementary grades compared to the alternative genotypes (ES=0.77).

Reading and Math Ability. A mixed linear model predicting reading ability at grades 1, 2, and 5, indicated that the interaction of *FKBP5* genotype with IPV was associated with the intercept, β=–4.199, SE=1.988, p=0.035, but not linear or quadratic growth. The finding indicates that carriers of two copies of the CATT haplotype who experienced IPV early in development exhibited lower levels of reading ability in the early elementary grades compared to the alternative genotypes. Carriers of two copies of the CATT haplotype who experienced IPV exhibited reading ability approximately a third of a standard deviation lower than carriers of two copies of the CATT haplotype not experiencing IPV (ES=0.39). There were no effects of the interaction of the CATT haplotype with IPV on mathematics ability at any time point.

Sensitivity Analyses

Adding maternal depressive symptoms at child age 24 months to the interaction model between CATT haplotype and IPV predicting cluster group assignment did not significantly improve model fit (p=0.800). Furthermore, the addition of maternal depression as a covariate did not change the findings, i.e., the FKBP5 x IPV remained a significantly predicted cluster order (p=0.036) while controlling for maternal depressive symptoms.

As an additional sensitivity check, we investigated the interaction between maternal CATT haplotype and IPV on child outcomes. As expected, the correlation between maternal and child genotype was high (r=0.530). We found that IPV did not significantly differ by either child genotype (p=0.410) or maternal genotype (p=0.300). These findings are consistent with previous studies where there is no correlation between maltreatment and genotype.

Through model comparisons, we found that an interaction between maternal CATT haplotype with IPV did not significantly improve the cluster order prediction (maternal CATT*IPV ~cluster order) to the model only including a main effect of IPV (IPV ~ cluster order) (p=0.25). Conversely, and as described in our original submission, the inclusion of the interaction between child CATT haplotype and IPV did significantly improve cluster order prediction (child CATT*IPV ~cluster order) compared to only the main effect of IPV (IPV ~ cluster order) (p=0.01). We also explored whether the IPV*child CATT haplotype predicting cluster order was further moderated by maternal CATT haplotype. This model (child CATT*IPV*maternal CATT ~ cluster order) did not improve in prediction accuracy compared to the reduced model (child CATT*IPV ~ cluster order).

.

Tables

TABLE S1. Number of participants with available data for each outcome

Variable	N	Mean	SD
Intimate partner violence mean 6–24 months	831	0.847	0.651
Cumulative risk mean 6–24 months*	891	0.011	0.639
Household chaos*	891	0.014	0.628
Fear-elicited emotional reactivity 6 months	785	0.106	0.210
Fear-elicited emotional reactivity 15 months	697	0.368	0.311
Fear-elicited emotional reactivity 24 months	663	0.369	0.344
Executive function 35 months	822	−0.550	0.547
Executive function 48 months	822	−0.127	0.515
Executive function 60 months	834	0.309	0.479
IQ 15 months	844	92.239	10.642
IQ 36 months	825	93.500	16.500
Teacher reported behavior problems grade 1	734	6.945	0.324
Teacher reported behavior problems grade 2	741	7.983	0.317
Teacher reported behavior problems grade 5	562	11.255	0.327
Reading ability W score grade 1	825	449.102	25.633
Reading ability W score grade 2	827	471.939	22.284
Reading ability W score grade 5	701	505.491	17.676
Math ability W score grade 1	825	455.004	20.212
Math ability W score grade 2	827	472.764	20.501
Math ability W score grade 5	701	507.029	21.731

TABLE S2. Demographic and characteristics of the sample by child *FKBP5* genotype

Variable	Total			Non-carriers N=432		One CATT Copy N=384		Two CATT Copies N=94	
	N		%	N	%	N	%	N	%
Sex									
Male	449		49.34	215	23.63	189	20.77	45	4.95
Female	461		50.66	217	23.85	195	21.43	49	5.38
Race									
African American	518		56.92	283	31.10	201	22.09	34	3.74
White	392		43.08	149	16.37	183	20.11	60	6.59
Maternal CATT haplotype[+]									
0 copies	382		47.63	278	34.66	103	12.84	1	0.12
1 copy	341		42.52	102	12.72	180	22.44	59	7.36
2 copies	79		9.85	0	0	59	7.36	20	2.49
	N	**Mean**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Intimate partner violence mean 6–24 months	831	0.847	0.651	0.857	0.605	0.853	0.629	0.912	0.714
Cumulative risk mean 6–24 months*	891	0.011	0.639	−0.032	0.664	0.059	0.586	0.059	0.663
Household chaos*	891	0.014	0.628	−0.030	0.624	0.047	0.612	0.087	0.655

[+]One-hundred and eight mothers did not provide saliva for genotyped or genotyping was not successful.
*The cumulative risk and household chaos scores have been transformed into z-scores

Figures

FIGURE S1. Schematic figure of the self-regulation model and how facets of the model influence emotional, behavioral and academic outcomes.

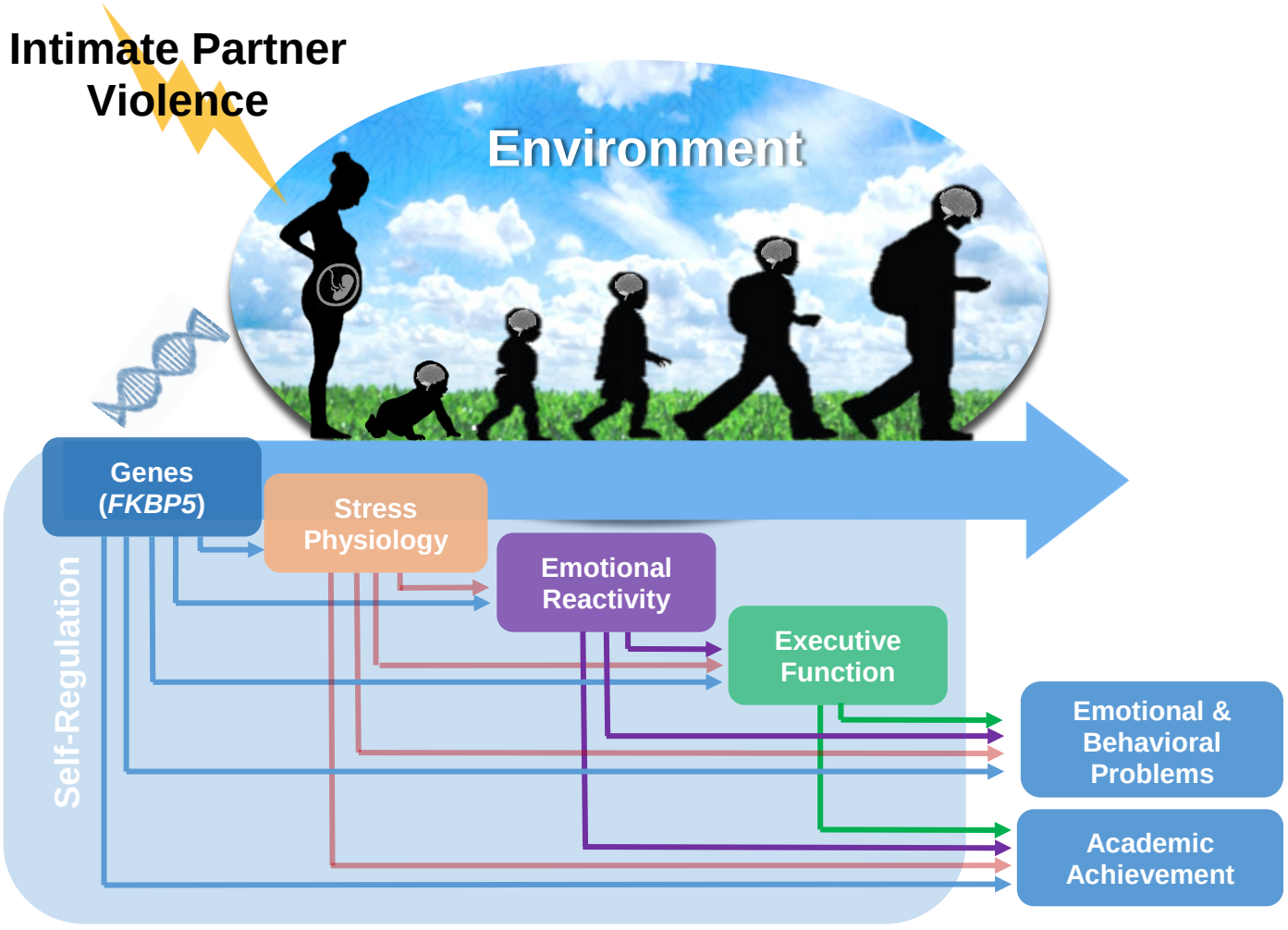


FIGURE S2. Cortisol levels prior to and following the stress paradigm at 7, 15 and 24 months broken down by CATT haplotype and exposure to intimate partner violence. The figure has been adjusted for the time of day the stress paradigm was administered, biological sex, cumulative risk, household chaos, state of residency, and ancestry markers.

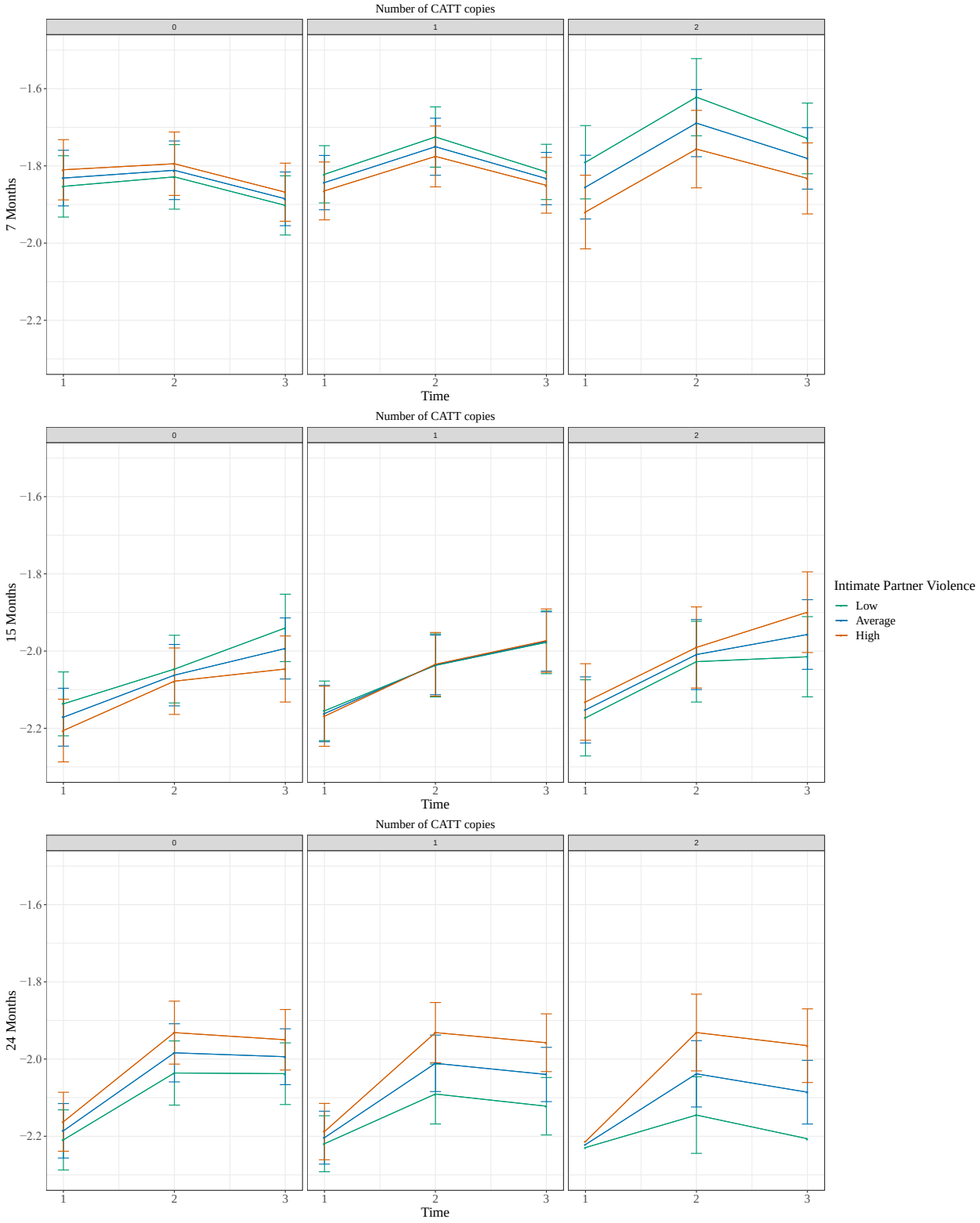


FIGURE S3. Pairwise linkage disequilibrium between genetic markers

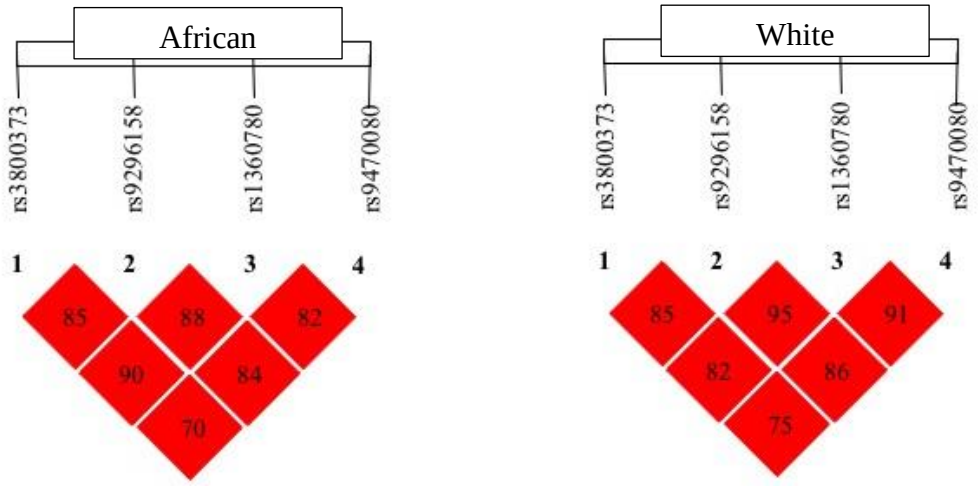


FIGURE S4. Flow chart of the statistical analyses conducted in the study

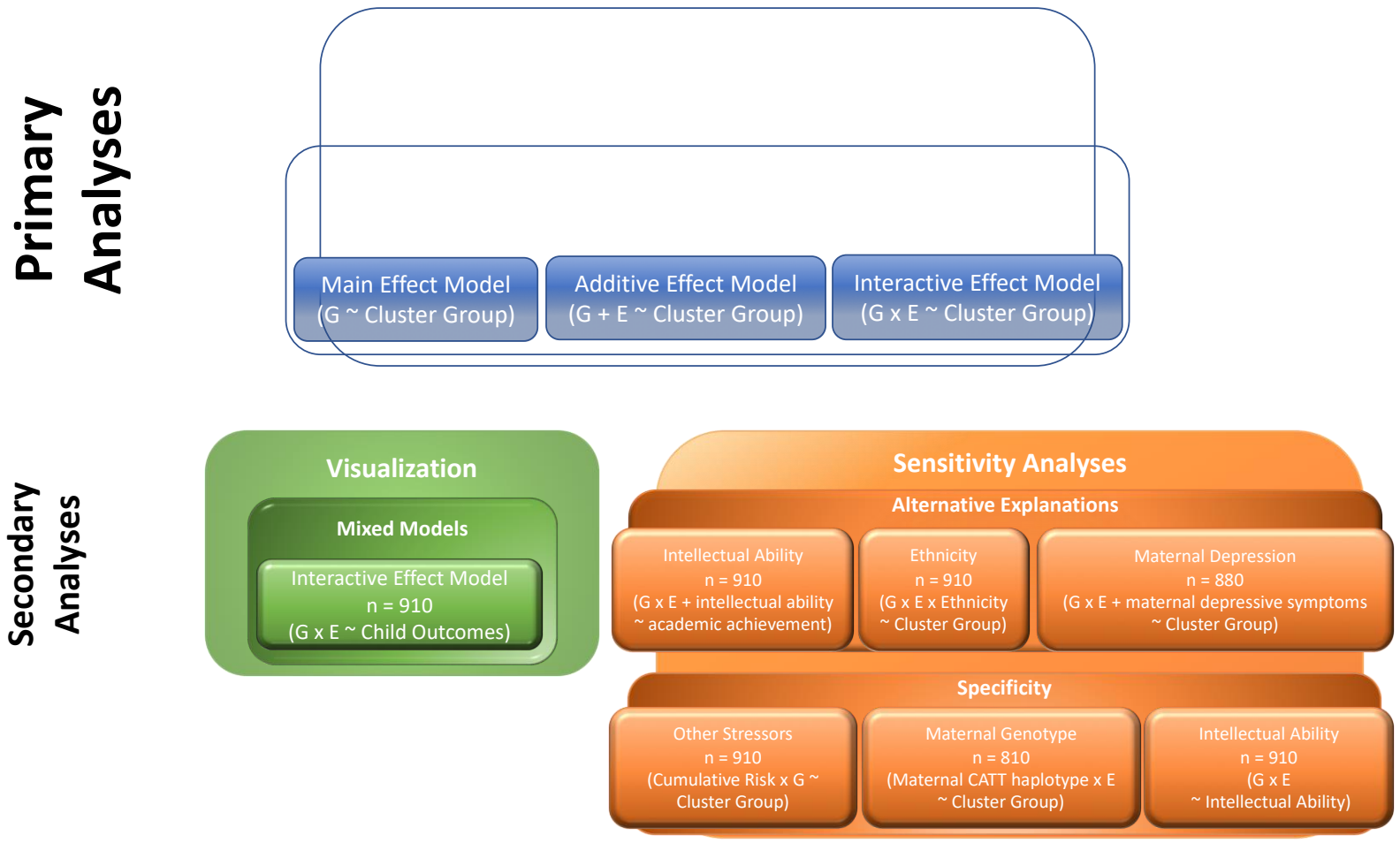
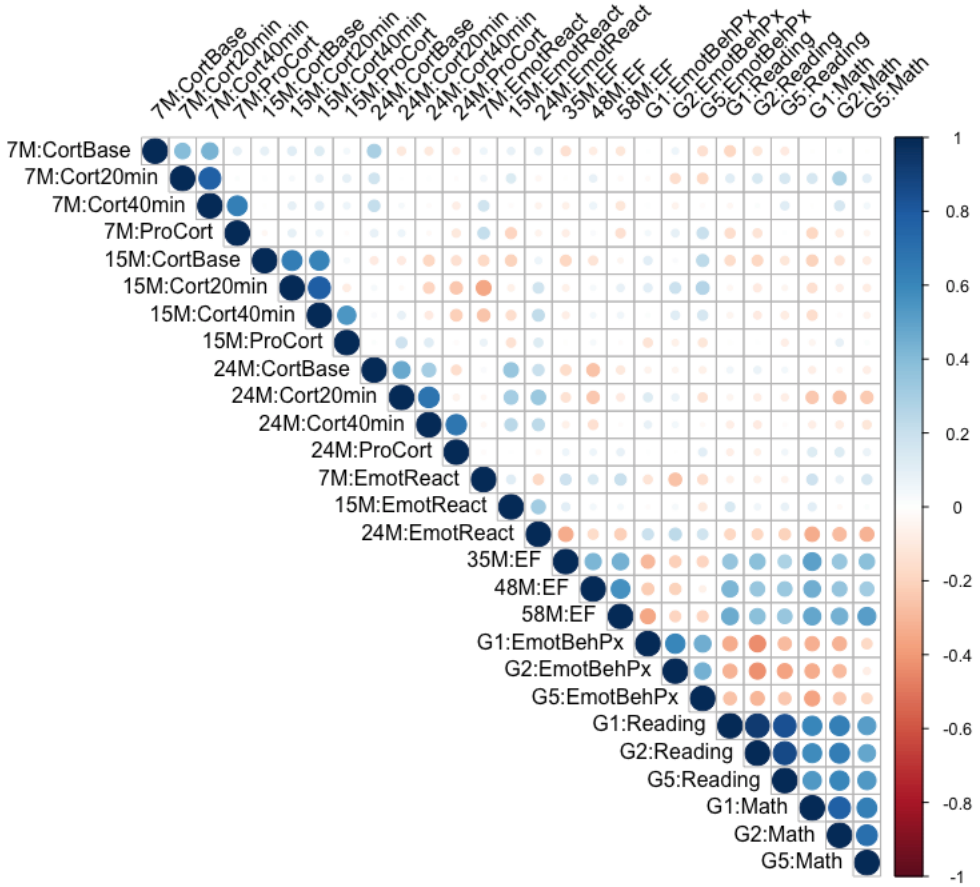


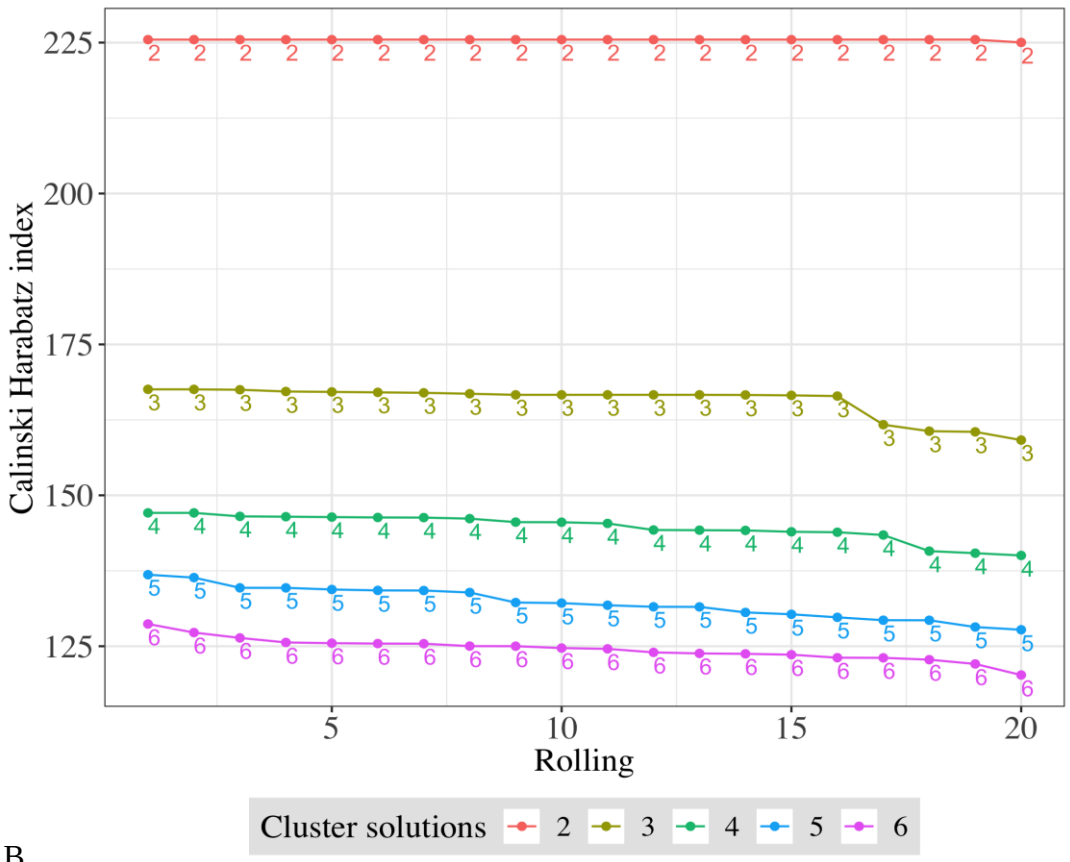
FIGURE S5. Correlation plot between outcome variables at each time point



Note. **CortBase=cortisol levels at baseline; Cort20min=cortisol levels 20 minutes after the stress paradigm; Cort40min=cortisol levels 40 minutes after the stress paradigm; ProCort=prolonged stress-induced cortisol reactivity; EmotReact=emotional reactivity; EF=executive function; EmotBehPx=emotional and behavioral problems.**

FIGURE S6. CH index for differing numbers of clusters in the (A) early development cluster analyses (i.e., prolonged cortisol and emotion reactivity) and (B) school-aged cluster analyses (i.e., emotional and behavioral difficulties and reading ability)

A



B

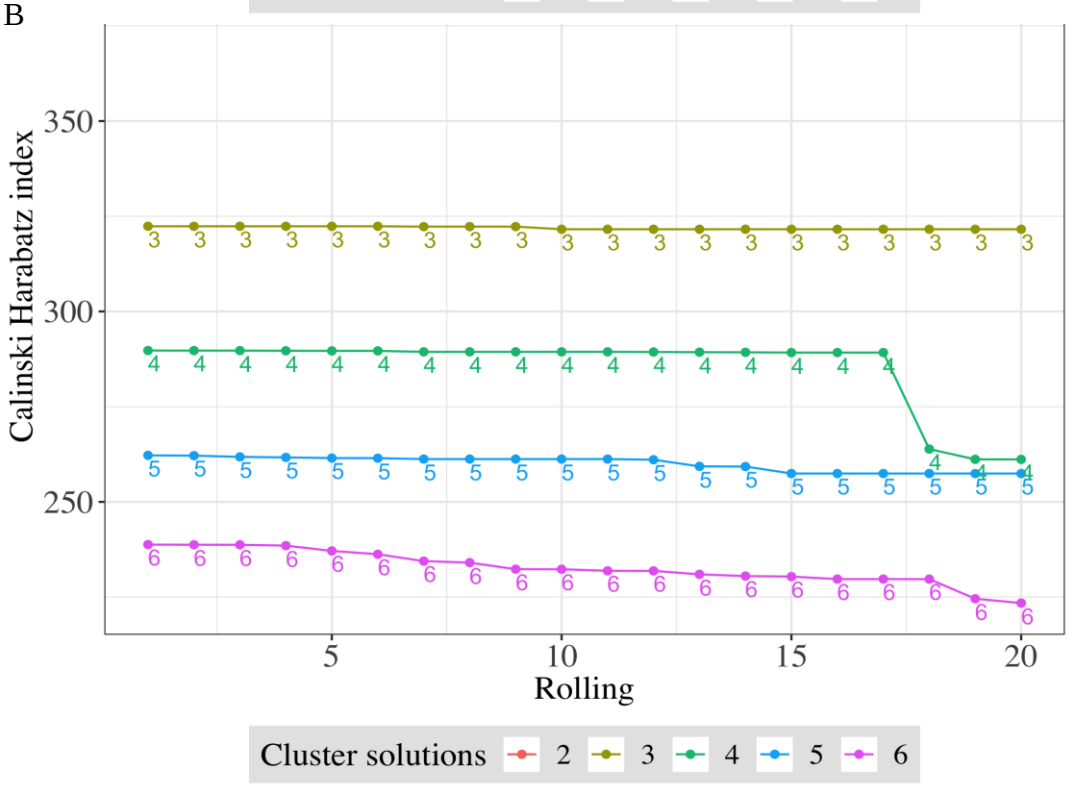
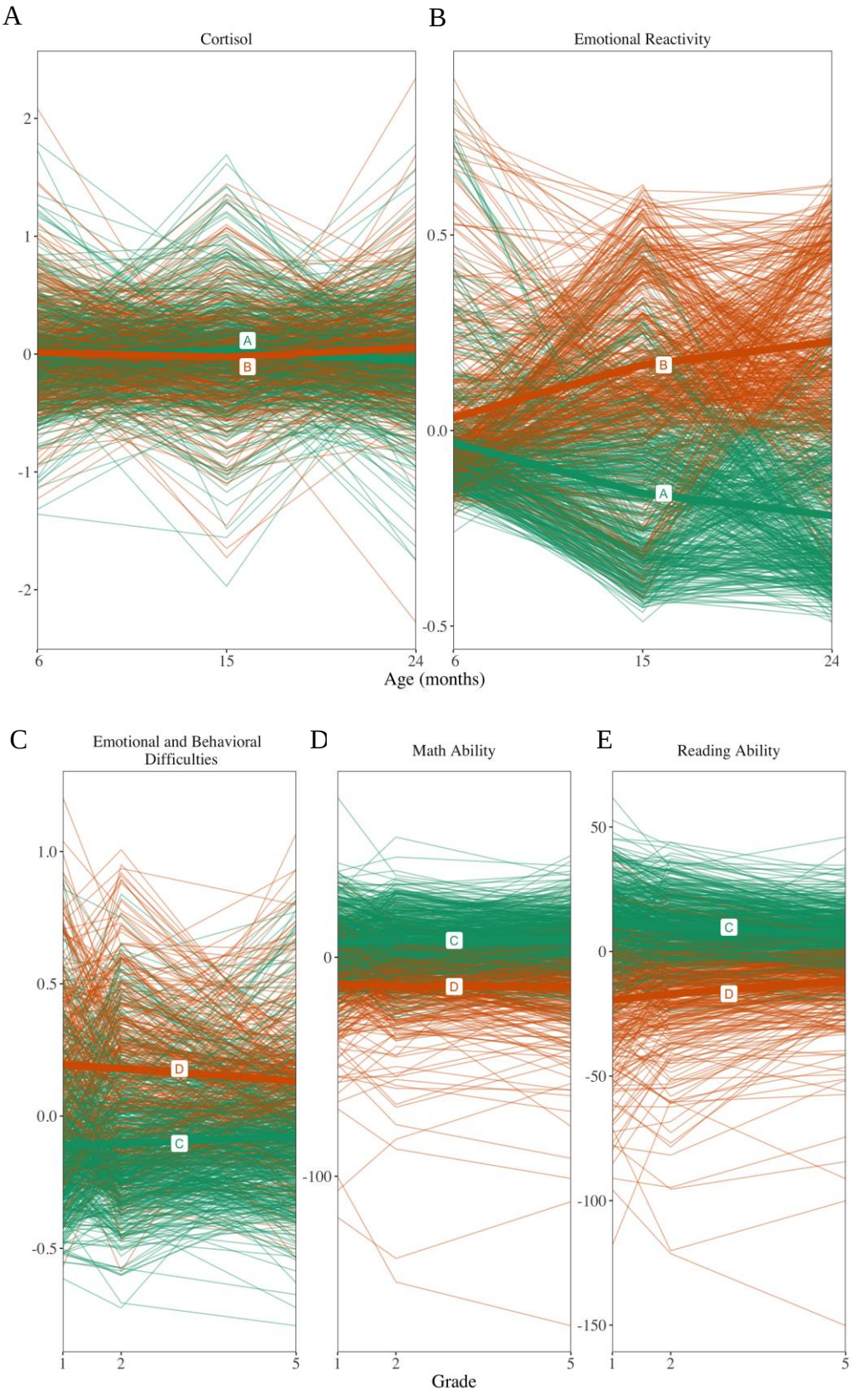


FIGURE S7. Trajectories of prolonged cortisol (A) and emotion reactivity (B) across time based on the two cluster solution from the early developmental cluster analyses and trajectories of emotional and behavioral difficulties (C), reading (D) and math ability (E) across time based on the two cluster solution from the school-aged cluster analyses



References

1. Buss KA, Goldsmith HH. Fear and anger regulation in infancy: Effects on the temporal dynamics of affective expression. Child development. 1998;69:359-374.
2. Kochanska G, Tjebkes JL, Fortnan DR. Children's emerging regulation of conduct: Restraint, compliance, and internalization from infancy to the second year. Child development. 1998;69:1378-1389.
3. Stifter CA, Braungart JM. The regulation of negative reactivity in infancy: Function and development. Developmental psychology. 1995;31:448.
4. Blair C, Granger DA, Kivlighan KT, Mills-Koonce R, Willoughby M, Greenberg MT, Hibel LC, Fortunato CK. Maternal and child contributions to cortisol response to emotional arousal in young children from low-income, rural communities. Developmental psychology. 2008;44:1095-1109.
5. Granger DA, Kivlighan KT, Fortunato C, Harmon AG, Hibel LC, Schwartz EB, Whembolua G-L. Integration of salivary biomarkers into developmental and behaviorally-oriented research: problems and solutions for collecting specimens. Physiology & behavior. 2007;92:583-590.
6. Harmon AG, Hibel LC, Rumyantseva O, Granger DA. Measuring salivary cortisol in studies of child development: watch out—what goes in may not come out of saliva collection devices. Developmental psychobiology. 2007;49:495-500.
7. Willoughby MT, Blair CB, Wirth RJ, Greenberg M. The measurement of executive function at age 5: psychometric properties and relationship to academic achievement. Psychological assessment. 2012;24:226.
8. Haberstick BC, Smolen A. Genotyping of three single nucleotide polymorphisms following whole genome preamplification of DNA collected from buccal cells. Behavior genetics. 2004;34:541-547.
9. Simmons TR, Flax JF, Azaro MA, Hayter JE, Justice LM, Petrill SA, Bassett AS, Tallal P, Brzustowicz LM, Bartlett CW. Increasing genotype-phenotype model determinism: application to bivariate reading/language traits and epistatic interactions in language-impaired families. Human heredity. 2010;70:232-244.
10. Hou L, Phillips C, Azaro M, Brzustowicz LM, Bartlett CW. Validation of a cost-efficient multi-purpose SNP panel for disease based research. PloS one. 2011;6:e19699.

.

11. Bruse SE, Moreau MP, Azaro MA, Zimmerman R, Hoffman A, Brzustowicz LM. Improvements to bead based oligonucleotide ligation SNP genotyping assays. Biotechniques. 2008;45:559.
12. Price AL, Butler J, Patterson N, Capelli C, Pascali VL, Scarnicci F, Ruiz-Linares A, Groop L, Saetta AA, Korkolopoulou P. Discerning the ancestry of European Americans in genetic association studies. PLoS genetics. 2008;4:e236.
13. Stekhoven DJ. missForest: Nonparametric missing value imputation using random forest. Astrophysics Source Code Library. 2015.
14. Genolini C, Alacoque X, Sentenac M, Arnaud C. kml and kml3d: R Packages to Cluster Longitudinal Data. Journal of Statistical Software. 2015;65:1-34.

.