Preregistration

*Last updated on Friday, March 10, 2023 at 06:46 PM*

## Overview

This preregistration document is based on ([Akker et al., 2021](#ref-akker2021)) for secondary data analyses. There are six parts, which you can jump to following the links below:

* [Part 1 - Study information](#part-1---study-information)
* [Part 2 - Data Description](#part-2---data-description)
* [Part 3 - Variables](#part-3---variables)
* [Part 4 - Knowledge of Data](#part-4---knowledge-of-data)
* [Part 5 - Analyses](#part-5---analyses)
* [Part 6 - Statement of Integrity](#part-6---statement-of-integrity)

## Part 1 - Study Information

### Q1: Title

“Within-person cognitive performance across abilities among adversity-exposed people in the SECCYD”

### Q2: Authors

* [Ethan S. Young](https://www.ethan-young.com/)1
* [Stefan Vermeent](https://www.stefan-vermeent.nl)1, 2
* [Willem E. Frankenhuis](http://www.willem.maartenfrankenhuis.nl/)1, 2
* [Marissa Nivison](https://icd.umn.edu/people/nivis004/)3
* [Jeffry A.Simpson](https://cla.umn.edu/about/directory/profile/simps108)3
* [Glenn I. Roisman](https://icd.umn.edu/people/roism001/)3

1Utrecht University

2Max Planck Institute for the Study of Crime, Security and Law

3University of Minnesota

### Q3: Research Questions

\*Note that some research questions/hypotheses use causal language. Although we are most interested in these causal questions, the data are observational. We retain the causal language here but will use non-causal language when relating any statistical test to these questions/hypotheses.

**R1:** How does adversity relate to overall performance compared to subtest performance on large cognitive test batteries? In other words, how does adversity relate to broad testing scores compared with more narrow subtest scores?

**R2:** How do impairment and enhancements in cognitive performance manifest across many cognitive domains?

**R3:** Do different adversity dimensions have similar or different patterns of effects across broad and narrow cognitive performance measures?

**R4:** Are the effects of adversity test-specific or test-general?

(Not sure that I want to keep R4 yet)

### Q4: Hypotheses

**H1:** We expect harshness and/or unpredictability to be associated with a lower overall within-person Woodcock Johnson (WJ) cognitive ability score (sum coded within-person intercept) .

**H2:** Compared with overall WJ scores, the effect of harshness and/or unpredictability will vary; some sub-tests will show lowered performance whereas others remain ‘intact’.

**H3:** If any sub-tests remain intact (or enhanced), they will be tests that depend less on formal crystallized knowledge and reading ability (i.e., short term memory, auditory processing, fluid intelligence).

## Part 2 - Data Description

### Q5: Dataset

We will use data from the National Institute of Child Health and Development (NICHD) Study of Early Childcare and Youth Development ([SECCYD](https://www.icpsr.umich.edu/web/ICPSR/series/00233), [Network, 2005](#Xc7c4e0c9d0e2771cb119bb9b72213bbaf258b7d)). The SECCYD was conducted in four waves across multiple sites and designed to be nationally representative, prospective, longitudinal study. The broad research goals of the study was to investigate the relation between childcare and development from infancy through adolescence and into early adulthood. Families were recruited for the NICHD SECCYD in 1991. A total of 1364 families met all the prescreening criteria. Below are detailed descriptions of each study phase:

* [Phase 1 (1991-1994)](https://www.icpsr.umich.edu/web/ICPSR/studies/21940)
* [Phase 2 (1995-1999)](https://www.icpsr.umich.edu/web/ICPSR/studies/21941)
* [Phase 3 (2000-2004)](https://www.icpsr.umich.edu/web/ICPSR/studies/21942)
* [Phase 4 (2005-2007)](https://www.icpsr.umich.edu/web/ICPSR/studies/22361)

### Q6: Public Availability

Data are publicly available. However, users must make an account with ISCPR (see [here](https://www.icpsr.umich.edu/rpxlogin)) anmd must sign a data use agreement. Users must also provide details about how data will be used (e.g., project description) and information (e.g., IRB approval). For each study phase, see more information about data access under “Access Restricted Data” tab.

### Q7: Data Access

Data can be accessed through the following links.

* [Phase 1 (1991-1994)](https://www.icpsr.umich.edu/web/ICPSR/studies/21940), doi: https://doi.org/10.3886/ICPSR21940.v6
* [Phase 2 (1995-1999)](https://www.icpsr.umich.edu/web/ICPSR/studies/21941), doi: https://doi.org/10.3886/ICPSR21941.v5
* [Phase 3 (2000-2004)](https://www.icpsr.umich.edu/web/ICPSR/studies/21942), doi: https://doi.org/10.3886/ICPSR21942.v6
* [Phase 4 (2005-2007)](https://www.icpsr.umich.edu/web/ICPSR/studies/22361), doi: https://doi.org/10.3886/ICPSR22361.v5

### Q8: Date of Download

* Ethan Young (lead author and data analyst)
  + Accessed data for the dependent variables on February 3rd, 2022
  + Accessed data for the independent variables on March 2nd, 2023
* Stefan Vermeent will not access the data
* Willem Frankenhuis will not access the data
* Marissa Nivison has access to the full dataset
* Jeffry Simpson will not access the data
* Glenn Roisman has access to the full data set

### Q9: Data Collection

Detailed information about recruitment, selection procedures, measures, and study methodology can be found [online](https://www.icpsr.umich.edu/web/ICPSR/series/00233).

### Q10: Codebooks

Detailed codebooks for each study wave can be downloaded at the following links:

* [Phase 1 (1991-1994)](https://www.icpsr.umich.edu/web/ICPSR/studies/21940/datadocumentation)
* [Phase 2 (1995-1999)](https://www.icpsr.umich.edu/web/ICPSR/studies/21941/datadocumentation)
* [Phase 3 (2000-2004)](https://www.icpsr.umich.edu/web/ICPSR/studies/21942/datadocumentation)
* [Phase 4 (2000-2004)](https://www.icpsr.umich.edu/web/ICPSR/studies/22361/datadocumentation)

Once variables for this study are selected, accessed, and ready for pre-processing/analysis, codebooks will be available [here](../codebooks/)

## Part 3 - Variables

### Q11: Manipulated Variables:

**Not applicable**

### Q12: Measured Variables

#### Covariates

* Gender
* Ethnicity
* Maternal education
  + 1 = less than high school
  + 2 = high school or general education diploma
  + 3 = some college or vocational degree
  + 4 = college degree
  + 5 = some graduate school or master’s degree
  + 6 = graduate degree greater than a master’s degree

#### Independent Variables

We are interested in two constructs: environmental harshness and unpredictability. However, in the literature, and in these data specifically, there have been different approaches to measuring them. There are also other, unexplored ways to capture them. Given the exploratory nature of our approach, we plan to compute both and explore their effects.

**Unpredictability, Past Approaches**

1. *Environmental Unpredictability*

This measure is based on Belsky et al. ([2012](#ref-belsky2012)) includes three variables that are standardized and averaged together over the relevant time period:

* Residential changes, or changes in address.
* Paternal transitions, or changes in father figures moving in or out of the home.
* Job changes, or changes in employment status for mothers and partners.

1. *Income Variation*

This measure is based on Li et al. ([2018](#ref-li2018)) computes the residual variance in income-to-needs ratios after a linear trend is fit to each participant.

**Harshness, Past Approaches**

Studies based on both the Belsky et al. ([2012](#ref-belsky2012)) and Li et al. ([2018](#ref-li2018)) approaches have used income-to-needs ratio. In most cases, an average score over the relevant period is the measure. However, some use an intercept after fitting a linear model to each person’s income data.

**The Current approach**

In addition to these classic measures, we plan to leverage data from the Census about the broader ecological context. Addresses were tracked for each participant over time. These addresses were geocoded and linked to the 1990 and 2000 decinniel Census blocks. Census blocks are the smallest geographical area that the Census measures. We will only use the 1990 Census blocks because all IVs in this study will be from when target participants were 0 - 54 months old.

The Census variables relevant here are the following:

* Percent of people living under the poverty line
* Median household income
* Gini coefficient, a metric for measuring income inequality
* Percent of unemployed individuals over 16 in the workforce
* Percent of occupied houses that are occupied by renters

We plan to compute mean and standard deviation scores for each Census measure over the addresses each participant lived at. For an overall neighborhood harshness score, we will standardize and avergae together each mean score. For an overall neighborhood change score, we will standardize and average together all standard deviation scores.

**Summary**

In short, we plan to analyze the following variables:

*Harshness*

* Census based neighborhood harhsness (from census variable averages)
* Average income-to-needs

*Unpredictability*

* Classic composite from Belsky et al. ([2012](#ref-belsky2012))
* Census-based neighborhood change (from standard deviations)
* Standard deviation of income-to-needs (this is a deviation from Li et al. ([2018](#ref-li2018)))

**Possible follow-ups**

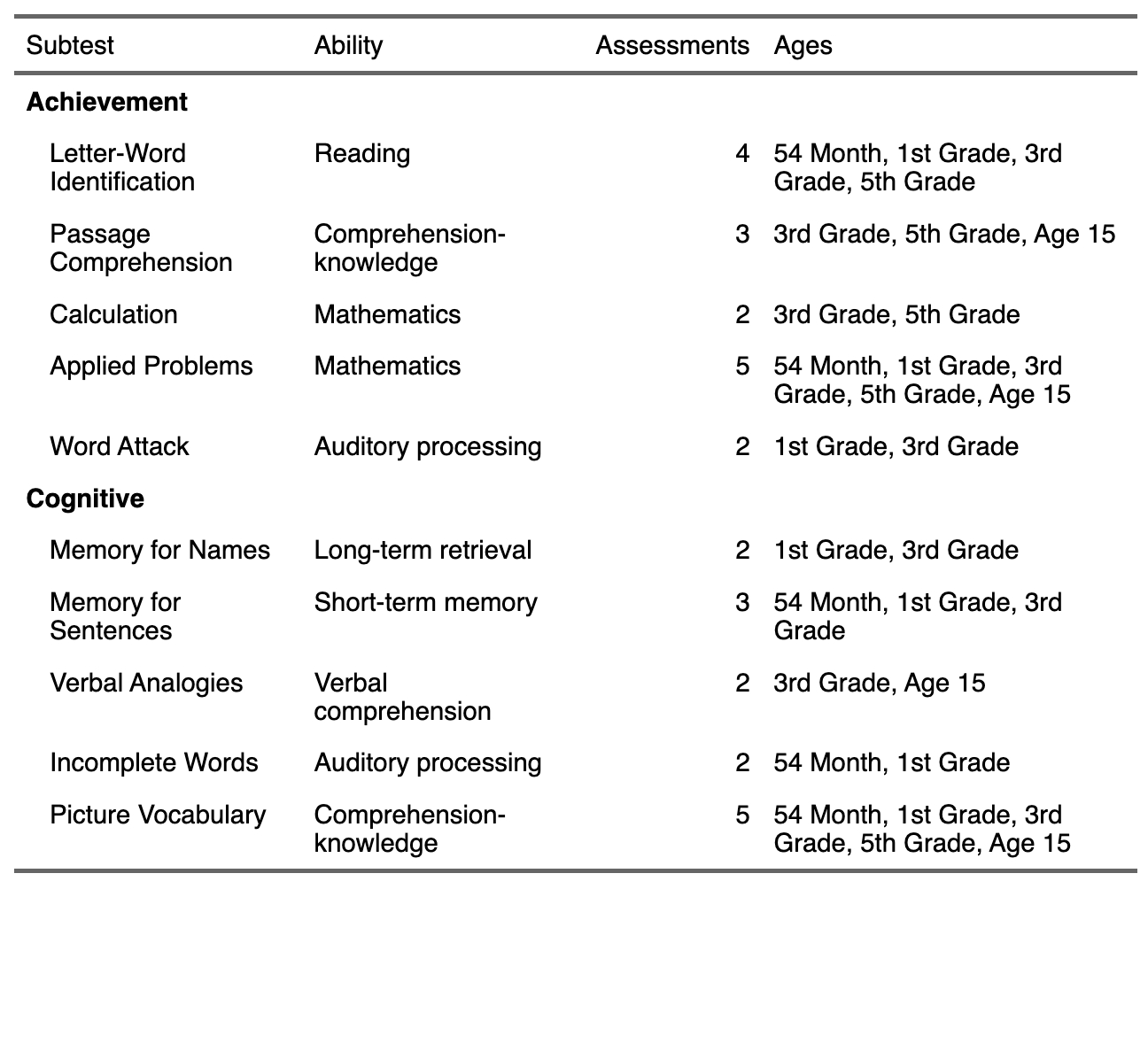
We might break apart composites. For exampe, we may analyze Census variables separately or items from the classic unpredictaiblity measure. If we do so, these analyses will be reported in a supplement, only. We may mention them in the main text, but they will not be interpreted to the same degree as the main analyses.

**Other possible (exploratory) variables:**

* Maternal depression
* Variability in maternal depression

#### Dependent Variables

The main dependent measures will come from the Woodcock-Johnson Cognitive and Achievement Tests ([Woodcock, 1990](#ref-woodcock1990); [Woodcock et al., 1990](#ref-woodcock1990a)).



**Scores**

For all tests, we will use standard scores. These scores are equivalent to IQ scores in that they use a mean of 100 and standard deviation of 15. This is useful when comparing many different tests.

**Aggregation strategy**

For each subtest, standard scores will be averaged over time to arrive at one score per subtest. For example, picture vocabulary was measured five times so overall picture vocabulary will be averaged over the five time points.

### Q13: Inclusion/Exclusion criteria

At the time of writing this preregistration, the only inclusion criteria are that participants should have a least one non-missing score on each subtest for at least one measurment period. This ensures that every case has at least one assessment of each subtest included in their overall average of all subtests. Participants must also have at least one score for the adversity measure. Each adversity measure will be analyzed in a separate model so the total samples size for each adversity measure may differ depending on missing data patterns (see Q14).

If it becomes clear that there are other inclusion/exclusion criteria, we will update the preregistration and/or report deviations in the the final manuscript. If there are many reasonable alternative criteria, we may use multiverse analysis to handle all combinations of reasonable and arbitrary inclusion/exclusion criteria.

### Q14: Missing data

*NOT DONE*

### Q15: Outliers

*NOT DONE*

### Q16: Sample Weights

**Not applicable**

## Part 4 - Knowledge of Data

### Q17: Relevant Publications

**No author has analyzed or worked with the Woodcock Johnson subtest scores prior to this preregistration.**

EY has not published any papers using this dataset. In 2014 and 2015, EY, JS, and GI submitted a paper to *Child Development* (rejected) and *Development and Psychopathology* (withdrawn). The paper used aggregated Woodcock Johnson scores over each assessment and used income-to-needs as a covariate.

MN has intimate knowledge of this dataset. MN has published two papers using the dataset. MN also has one manuscript in press and one under review using the data. MN has not analyzed or used the subtest variables in the current preregistration.

GI has intimate knowledge of this dataset. GI is a co-principle investigator on the project and has published many papers using the data. Variables analyzed in GI publications relate mostly to the dependent variables in this project. Those relevant to the current preregistration include:

* Bleil et al. ([2021](#ref-bleil2021))
* Cottrell et al. ([2015](#ref-cottrell2015))
* Monti et al. ([2014](#ref-monti2014))
* Fraley et al. ([2013](#ref-fraley2013))
* Roisman et al. ([2012](#ref-roisman2012))
* Burt & Roisman ([2010](#ref-burt2010))

JS has also worked with the current dataset. Variables in JS publications relate mostly to the independent variables used in this study. Those relevant to the current preregistration include:

* Hartman et al. ([2018](#ref-hartman2018a))
* Sung et al. ([2016](#ref-sung2016))

SV and WF have no prior experience with the data.

### Q18: Prior Knowledge

From prior work, we have some knowledge of how income, maternal education, and quality of maternal caregiving (observations of maternal sensitivity) are associated with aggregated Woodcock Johnson scores. For example, Fraley et al. ([2013](#ref-fraley2013)) find correlations between Woodcock Johnson composites over 5 assessment periods and income-to-needs (averaged over 6, 15, 24, and 36 months; *rs* range = .34-.37), maternal education (*rs* range = .41 - .47), and maternal sensitivity (*rs* range = .40 - .47) . These findings give us a strong prior that harshness/poverty will be associated with a lower within person average Woodcock Johnson score.

## Part 5 - Analyses

### Q19: Hypotheses -> Statistical Tests

Below is a conceptual depiction of our analyses:

|  |
| --- |
| Figure 1: We are interested in the effect of each adversity measure on a person’s overall score, measured as a formative average of each subtest. (A) is the main effect of adversity on overall performance. (B) is the main effect of a subtest. (C) is the simple effect (slope) of adversity for a particular subtest. (D) is the interaction effect that measures the difference between A and C. A significant simple slope means the C ≠ 0 and a significant interaction means A ≠ C. So, when C is significant, it means that adversity affects performance. When D is significant, it means that adversity affects a subtest in a different way than A (overall pattern). |

We will use a mixed effects linear regression to test Hypotheses 1 & 2 using the lmer. To do so, we proceed in three steps:

1. We will standardize (z-score) our independent variables. This centers the IV at 0 and scales them with standard deviation = 1.

example\_data1 <-   
 seccyd\_dvs\_wj\_data2 |>   
 group\_by(id) |>   
 summarize(  
 # Average scores over time to single scores for each subtest  
 across(starts\_with("wj\_"), list(mean = ~ mean(.x, na.rm = T), n = ~sum(!is.na(.x))))  
 ) |>   
 mutate(across(everything(), ~ifelse(is.nan(.x), NA, .x))) |>   
 select(id, ends\_with("mean")) |>   
 mutate(  
 id = 1:n(),  
 # Make up an adversity variable for the preregistration and standardize (step 1)  
 adversity = rnorm(n()) |> scale() |> as.numeric()  
 )

1. We then stack the data into ‘long’ format. There are 10 subtests, so each participant will have 10 associated rows. Two new columns are created, one indicating the subtest type and one indicating each participant’s associated score.

Below is an example of one participant’s data:

example\_data2 <-   
 example\_data1 |>   
 # Stack data so that eadh participant has one row per subtest score (step 2)  
 pivot\_longer(c(-id, -adversity), names\_to = "wj\_sub\_test", values\_to = "score")  
  
# Show an example of one participants data structure  
example\_data2 |>   
 filter(id == 1) |>   
 knitr::kable()

| id | adversity | wj\_sub\_test | score |
| --- | --- | --- | --- |
| 1 | 1.20877 | wj\_picvo\_mean | 108.7500 |
| 1 | 1.20877 | wj\_vrba\_mean | 139.5000 |
| 1 | 1.20877 | wj\_pscmp\_mean | 133.0000 |
| 1 | 1.20877 | wj\_appld\_mean | 130.5000 |
| 1 | 1.20877 | wj\_memse\_mean | 115.0000 |
| 1 | 1.20877 | wj\_incom\_mean | 128.0000 |
| 1 | 1.20877 | wj\_memna\_mean | 117.0000 |
| 1 | 1.20877 | wj\_lwid\_mean | 138.6667 |
| 1 | 1.20877 | wj\_wrdat\_mean | 134.0000 |
| 1 | 1.20877 | wj\_calc\_mean | 143.0000 |

1. Next, we apply a sum coded contrast to the subtest index column. This means the intercept in the mixed effect model reflects the grand mean of all WJ subtest scores. Sum coding allows us to compare the effect of adversity on each subtest to the mean of all tests.

For example, the underlying contrasts would look like the following:

example\_data3 <-   
 example\_data2 |>   
 mutate(  
 # Apply sum contrasts (step 3)  
 wj\_sub\_test = faux::contr\_code\_sum(wj\_sub\_test)  
 )  
  
# Show the contrast scheme for the mixed model  
example\_data3 |>   
 pull(wj\_sub\_test) |>   
 attr("contrasts") |>   
 as.data.frame() |>   
 rownames\_to\_column(var = "test") |>   
 mutate(  
 test = str\_replace\_all(test,"^wj\_(.\*)\_mean$", "\\1")  
 ) |>   
 rename\_with(  
 .cols = everything(),   
 ~str\_replace\_all(.x ,"^.wj\_(.\*)\_mean-intercept", "\\1")  
 ) |>   
 knitr::kable()

| test | appld | calc | incom | lwid | memna | memse | picvo | pscmp | vrba |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| appld | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| calc | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| incom | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| lwid | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| memna | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| memse | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| picvo | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| pscmp | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| vrba | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| wrdat | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |

1. We fit a linear mixed effects model with the following terms
   * contrast coded subtest
   * adversity (standardized)
   * interaction between contrast coded subtest and adversity
   * random intercept for participants

# Fit model  
subtest\_model <- lmer(score ~ wj\_sub\_test \* adversity + (1|id), data = example\_data3)

Below is an example of the output from the above analysis:

**Unstandardized Parameters**

# Standardized and Unstandardized parameters  
subtest\_model |>   
 parameters::parameters() |>   
 mutate(across(where(is.numeric), ~round(.x, 3))) |>   
 mutate(  
 Parameter = str\_replace\_all(Parameter, "^wj\_sub\_test.wj\_(.\*)\_mean-intercept","\\1"),  
 Parameter = str\_replace(Parameter, ":", " \* ")  
 ) |>   
 filter(Effects == "fixed") |>   
 select(Parameter, Coefficient, SE, p) |>   
 knitr::kable()

| Parameter | Coefficient | SE | p |
| --- | --- | --- | --- |
| (Intercept) | 105.338 | 0.321 | 0.000 |
| appld | 2.566 | 0.259 | 0.000 |
| calc | 6.541 | 0.268 | 0.000 |
| incom | -9.193 | 0.266 | 0.000 |
| lwid | 1.600 | 0.259 | 0.000 |
| memna | -0.233 | 0.264 | 0.378 |
| memse | -8.311 | 0.260 | 0.000 |
| picvo | -2.697 | 0.259 | 0.000 |
| pscmp | 3.071 | 0.267 | 0.000 |
| vrba | 5.443 | 0.269 | 0.000 |
| adversity | 0.302 | 0.322 | 0.348 |
| appld \* adversity | -0.059 | 0.259 | 0.821 |
| calc \* adversity | -0.217 | 0.270 | 0.422 |
| incom \* adversity | -0.349 | 0.267 | 0.191 |
| lwid \* adversity | -0.072 | 0.259 | 0.781 |
| memna \* adversity | 0.343 | 0.266 | 0.198 |
| memse \* adversity | 0.015 | 0.261 | 0.954 |
| picvo \* adversity | 0.308 | 0.259 | 0.235 |
| pscmp \* adversity | -0.213 | 0.269 | 0.430 |
| vrba \* adversity | 0.023 | 0.272 | 0.931 |

**Standardized Parameters**

# Standardized and Unstandardized parameters  
subtest\_model |>   
 parameters::standardize\_parameters() |>  
 mutate(across(where(is.numeric), ~round(.x, 3))) |>   
 mutate(  
 Parameter = str\_replace\_all(Parameter, "^wj\_sub\_test.wj\_(.\*)\_mean-intercept","\\1"),  
 Parameter = str\_replace(Parameter, ":", " \* ")  
 ) |>   
 select(-CI) |>   
 knitr::kable()

| Parameter | Std\_Coefficient | CI\_low | CI\_high |
| --- | --- | --- | --- |
| (Intercept) | -0.005 | -0.048 | 0.037 |
| appld | 0.173 | 0.139 | 0.207 |
| calc | 0.441 | 0.406 | 0.476 |
| incom | -0.620 | -0.655 | -0.585 |
| lwid | 0.108 | 0.074 | 0.142 |
| memna | -0.016 | -0.050 | 0.019 |
| memse | -0.560 | -0.595 | -0.526 |
| picvo | -0.182 | -0.216 | -0.148 |
| pscmp | 0.207 | 0.172 | 0.242 |
| vrba | 0.367 | 0.331 | 0.403 |
| adversity | 0.020 | -0.022 | 0.063 |
| appld \* adversity | -0.004 | -0.038 | 0.030 |
| calc \* adversity | -0.015 | -0.050 | 0.021 |
| incom \* adversity | -0.023 | -0.059 | 0.012 |
| lwid \* adversity | -0.005 | -0.039 | 0.029 |
| memna \* adversity | 0.023 | -0.012 | 0.058 |
| memse \* adversity | 0.001 | -0.033 | 0.035 |
| picvo \* adversity | 0.021 | -0.013 | 0.055 |
| pscmp \* adversity | -0.014 | -0.050 | 0.021 |
| vrba \* adversity | 0.002 | -0.034 | 0.037 |

### Q20: Predicted effect sizes

We do not have specific predictions for effect sizes but we deem standardized regression coefficients for interaction effects = .10 (or higher) and -.10 (or lower) meaningful. These effects would indicate that an effect of adversity on a subtest score is meaningfully more negative or more positive.

Effects between -.10 and .10 are not of interest for determining differences between subtests and the overall average. However, we are interested in determining if simple effects falling between -.10 and .10 are consitent with an effect of 0. For this scenario, we will use equivalence testing with -.10 and .10 as bounds.

### Q21: Statistical Power

See our full power analysis [here](power-analysis/) and the code [here](../scripts/power-simulation.R).

In short, we used a simulation approach to conduct power analyses. We simulated adversity scores but used actual Woodcock Johnson test scores. Simulations revealed, that, with a sample size of (*N* = 1156), the smallest interaction effect we can detect is = -.075 (or .075) with 90% power, if error is small. When error is larger, we can detect the same effect size with only 65% power. However, even with larger error, we can detect a = -.10 (or .10) with 83% power.

### Q22: Inferential Criteria

Our inferential criteria will use p < .05 for interaction effects between adversity and subtest type and for simple effects of adversity for each subtest.

We interested in three types of effect sizes:

1. The main effect of each adversity measure (tested in separate models).
2. The simple effect of adversity for each subtest.
3. The interaction effect between an adversity measure and subtest.

We are also interested in null effects. Based on prior work, we expect the main effect of adversity to be negative. However, if main effects were zero for any adversity measure, we would be interested in determining if the effect *is* zero by using equivalence testing ([Lakens et al., 2018](#ref-lakens2018)).

For simple effects (effect size 2), we will use simple slopes analysis. These tests determine whether the simple effect is different from zero. For these effects, we are interested if they are indeed 0 using equivalence testing. We are also interested positive slopes of any size (that are not equivalent to zero).

For negative slopes, we are interested in effects that are different from the overall main effect of adversity. This is where the interaction term is important. The interaction effect tests whether the effect of adversity for a given subtest is different from the main effect of adversity on overall performance. We are interested in both more negative and more positive than expected slopes compared to the main effect.

Based on our minimum effect size of interest criterion and our power analysis, we are interested in interaction effects where = -.10 (or lower) or = .10 (or higher). We will use these bounds to inform our equivalence tests, where effects are null, to determine if they are consistent with an effect = 0.

### Q23: Assumption Violations

If any assumptions are violated, we will update the preregistration and/or report deviations from the preregistration. However, we do not anticipate any serious violations.

### Q24: Evaluating Strength, Reliability, and Robustness

If there are arbitrary data processing decisions, we may use multiverse analysis to systematically explore their effect.

### Q25: Exploratory Analyses

We may or may not do serveral exploratory analyses listed below:

1. Fit structural equation models to Woodcock Johnson data to examine how adversity relates to an overall latent variable and to residual test variance in each subtest
2. Depending on the main results, we may also look at subtest scores over time to examine developmental trajectories in those tests.
3. We may run additional analyses identical to our main analyses but with different adversity measures that are not central to our framework.

## Part 6 - Statement of Integrity

The authors of this preregistration state that they filled out this preregistration to the best of their knowledge and that no other preregistration exists pertaining to the same hypotheses and dataset.

## References

Akker, O. R. van den, Weston, S., Campbell, L., Chopik, B., Damian, R., Davis-Kean, P., Hall, A., Kosie, J., Kruse, E., Olsen, J., Ritchie, S., Valentine, K. D., Veer, A. van ’t, & Bakker, M. (2021). Preregistration of secondary data analysis: A template and tutorial. *Meta-Psychology*, *5*. <https://doi.org/10.15626/MP.2020.2625>

Belsky, J., Schlomer, G. L., & Ellis, B. J. (2012). Beyond cumulative risk: Distinguishing harshness and unpredictability as determinants of parenting and early life history strategy. *Developmental Psychology*, *48*(3), 662–673. <https://doi.org/b7r3m4>

Bleil, M. E., Spieker, S. J., Gregorich, S. E., Thomas, A. S., Hiatt, R. A., Appelhans, B. M., Roisman, G. I., & Booth-LaForce, C. (2021). Early life adversity and pubertal timing: Implications for cardiometabolic health. *Journal of Pediatric Psychology*, *46*(1), 36–48.

Burt, K. B., & Roisman, G. I. (2010). Competence and psychopathology: Cascade effects in the NICHD Study of Early Child Care and Youth Development. *Development and Psychopathology*, *22*(3), 557–567. <https://doi.org/bc45zc>

Cottrell, J. M., Newman, D. A., & Roisman, G. I. (2015). Explaining the blackwhite gap in cognitive test scores: Toward a theory of adverse impact. *Journal of Applied Psychology*, *100*, 1713–1736. <https://doi.org/f7zb56>

Fraley, R. C., Roisman, G. I., & Haltigan, J. D. (2013). The legacy of early experiences in development: Formalizing alternative models of how early experiences are carried forward over time. *Developmental Psychology*, *49*(1), 109.

Hartman, S., Sung, S., Simpson, J. A., Schlomer, G. L., & Belsky, J. (2018). Decomposing environmental unpredictability in forecasting adolescent and young adult development: A two-sample study. *Development and Psychopathology*, *30*(4), 1321–1332. <https://doi.org/gjqzzj>

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. <https://doi.org/10.1177/2515245918770963>

Li, Z., Liu, S., Hartman, S., & Belsky, J. (2018). Interactive Effects of Early-Life Income Harshness and Unpredictability on Children’s Socioemotional and Academic Functioning in Kindergarten and Adolescence. *Developmental Psychology*, *54*(11), 2101–2112. <https://doi.org/gfmd6w>

Monti, J. D., Pomerantz, E. M., & Roisman, G. I. (2014). Can parents’ involvement in children’s education offset the effects of early insensitivity on academic functioning? *Journal of Educational Psychology*, *106*, 859–869. <https://doi.org/gm2bz2>

Network, N. E. C. C. R. (2005). *Child care and child development: Results from the NICHD study of early child care and youth development* (pp. xx, 474). The Guilford Press.

Roisman, G. I., Newman, D. A., Fraley, R. C., Haltigan, J. D., Groh, A. M., & Haydon, K. C. (2012). Distinguishing differential susceptibility from diathesisstress: Recommendations for evaluating interaction effects. *Development and Psychopathology*, *24*(02), 389–409. <https://doi.org/gd59f7>

Sung, S., Simpson, J. A., Griskevicius, V., Kuo, S. I.-C., Schlomer, G. L., & Belsky, J. (2016). Secure infant-mother attachment buffers the effect of early-life stress on age of menarche. *Psychological Science*, *27*(5), 667–674. <https://doi.org/f8n843>

Woodcock, R. W. (1990). Theoretical Foundations of the Wj-R Measures of Cognitive Ability. *Journal of Psychoeducational Assessment*, *8*(3), 231–258. <https://doi.org/ft7mjn>

Woodcock, R. W., Johnson, M. B., & Mather, N. (1990). *Woodcock-Johnson psycho-educational battery– revised*. DLM Teaching Resources.