# Abstract

The idea that some abilities might be enhanced by adversity is gaining traction. For example, research leveraging the hidden talents approach has uncovered a few specific abilities enhanced by exposure to particular forms of adversity in a given context. Yet, in order for a field to grow, we must not dig too deep, too fast. In this paper, we compliment adaptation-based research with principled exploration. To do so, we draw on the basic insights of adaptation-based research: 1) enhanced performance manifests within individuals and 2) reduced and enhanced performance can co-occur. Although commonly assumed, these assertions are rarely tested. To do so, a variety of ability measures are needed that examine relative performance differences. However, rather than using adaptive-logic to predict which abilities are enhanced or reduced, we develop statistical criteria to help interpret three different data patterns: reduced, enhanced, and intact performance. We use these criteria to analyze data from the Study of Early Childcare and Youth Development (SECCYD) to examine how adversity shapes within-person performance across 10 abilities in the Woodcock Johnson Cognitive and Achievement test battery. Our goals are to document adversity-shaped cognitive profiles, identify possible drivers of reduced overall performance, map out sets of ‘intact’ abilities, and discover new enhanced abilities. We argue that principled exploration with clear criteria can help break new ground, re-map old territory, and fuel theory development. Our approach thus offers a valuable complement to the adaptive-logic approach that has dominated this emerging area of research to date.

# How does adversity relate to performance across different abilities in the same person?

Developmental science commonly asserts that adversity-exposure during development reduces cognitive performance, a claim founded on decades of empirical findings (Duncan et al., 2017; Farah et al., 2006; Hackman et al., 2010; McLaughlin et al., 2019; Raby et al., 2015). In recent years, however, adaptation-based frameworks, rooted in the idea that adversity might enhance certain abilities, have complemented this work—and it is gaining traction (Ellis et al., 2017; Ellis et al., 2020; Frankenhuis, Young, et al., 2020; Frankenhuis & de Weerth, 2013; Frankenhuis & Nettle, 2020). Since its inception, the goal of adaptation-based frameworks has been to inspire a more well-rounded view of adversity and its influence on abilities—one that incorporates both the struggles and strengths of people from disadvantaged backgrounds (Frankenhuis & de Weerth, 2013). As it develops, the core task of adaptation-based research is to “uncover a high-resolution map of specific cognitive abilities that are enhanced as a result of growing up under high-adversity conditions” (Ellis et al., 2017, p. 562). To uncover this map, researchers have used confirmatory study designs, which has gleaned useful insights. Yet, to cultivate growth in an emerging research program—where there is little known and much to learn—we must not dig too deep, too fast. Without complimentary approaches, exclusive use of confirmatory designs can create tunnel vision and miss new insights (McIntosh, 2017; Roisman, 2021; Rozin, 2001; Scheel et al., 2021).

In this paper, we use a complimentary approach to confirmatory research: principled exploration. To guide our exploration, we leverage on two basic insights from adaptation-based research: 1) enhanced performance manifests within individuals and 2) reduced and enhanced performance can co-occur. The first insignt implies we need designs and models that can tease apart both within- and between-person peformance differences. The second suggests that–to map a broader set of adveristy-related abiities–we must examine many more abilities measured within the same person. Doing so will allow us to capture cogntiive performance profiles. These profiles comprise three types of data patterns: reduced, intact, and enhanced performance. Past research has focused on reduced and enhanced performance on specific ability tests. However, we know little about intact abilities, or cases where performance is unrelated to adversity exposure. Thus, our goal is to document adversity-shaped cognitive profiles that include reduced, ‘intact’ abilities, and enhanced test performance patterns.

# **Essential Features and Empirical Insights from Adaptation-based Frameworks**

Adaptation-based research has several essential features. First, it assumes development shapes the individual, and their abilities, to fit the local environment (Frankenhuis, Young, et al., 2020). Second, because environments differ in the challenges they pose (resource-scarcity versus violence exposure), development shapes abilities according to specific challenges. Thus, one’s abilities are thought to match the challenges of one’s lived experience. These features are useful guideposts for confirmatory hypothesis generation. Using them as building blocks, it is easy to construct an intuitive bridge between an ability and an environmental challenge. For example, a researcher might identify a specific challenge posed by a dimension of adversity (e.g., threats to safety in high-crime neighborhoods) and an ability needed to meet the challenge (e.g., enhanced threat detection).

This approach is appealing because it forces researchers to be specific and logically tie together challenges and abilities. It has also been successful in discovering a handful of interesting adversity-enhanced abilities, especially in harsh and unpredictable environments. For example, past work has proposed that constantly changing environments (i.e., unpredictable environments) might shape the ability to track and respond to changing information. Using this logic, research build an intuitive bridge between changing environments and two abilities–attention-shifting and working memory updating—and some empirical data are consistent with this logic (Fields et al., 2021; Mittal et al., 2015; Nweze et al., 2021; Young et al., 2018). However, there are two limitations to this approach. First, previous studies are difficult to compare because they use different measures and designs. Second, the logic behind confirmatory hypotheses is easily flipped. For example, exposure to unpredictable environments is thought to reduce inhibition, or the ability to resist distractions. If threats and opportunities arise, it is important to quickly respond, rather than ignore them to focus on a long-term goal. But we can assert the exact opposite. For example, inhibition might be enhanced by unpredictable environments because it helps to focus on what is important when there are constant distractions.

Adaptation-based research has also focused on testing content, or the notion that performance should improve when the testing content matches the lived experience of people exposed to adversity. For example, studies have examined relational memory, attention shifting, and working memory task performance using more ecologically relevant content (e.g., social dominance, real-world, and socioemotional stimuli) compared to neutral or abstract content. In some cases, ecologically relevant content appeared to equalize performance for people exposed to adversity, but this depends on the specific adversity measure and task (Frankenhuis, de Vries, et al., 2020; Rifkin-Graboi et al., 2021; Young et al., 2022). Yet, in other studies, conditions thought to be well-matched to the lived experience of those exposed to adversity actually lower performance. For example, youth from poverty tended to score lower on math items about social relations, money, and food—items thought to be particularly relevant to lived experience—compared to other math items (Duquennois, 2022; Muskens, 2019).

In light of various caveats, this body of work has generated at least two general insights. First, although it is possible for adversity to enhance performance between individuals (e.g., low versus high adversity exposure), empirical findings suggest effects mostly occur within individuals (Fields et al., 2021; Frankenhuis, de Vries, et al., 2020; Young et al., 2022). Second, enhanced performance appears to be highly context specific—enhancements depend on testing content, context, and ability type (Fields et al., 2021; Frankenhuis, de Vries, et al., 2020; Mittal et al., 2015; Nweze et al., 2021; Young et al., 2018; Young et al., 2022). Yet, adaptation-based studies have looked for abilities in an isolated and piecemeal fashion, in part, because confirmatory designs tend to narrow a study’s scope. This means we know little about enhanced abilities compared with the broad landscape of ability measures.

# **Motivating Principled Exploration**

We believe that adaptation-based frameworks can provide useful guideposts, but one should use shovels, not scalpels, when breaking new ground. Emerging research programs have yet to lay basic groundwork for testing theories, such as auxiliary assumptions or boundary conditions Scheel et al. (2021)]. Our aim is to complement adaptation-based, confirmatory research with principled exploration (Flournoy et al., 2020; Rozin, 2001). We see two benefits of this approach. The first is to re-examine established patterns with a new lens. For example, both deficit- and adaptation-based perspectives assume that adversity should reduce performance on standard assessments of cognitive ability (Ellis et al., 2020; Frankenhuis, Young, et al., 2020; Hackman et al., 2010; McLaughlin et al., 2019; Ursache & Noble, 2016). Yet, these tests are often comprised of many different subtests, and may show unique patterns that diverge from widely used composite scores. The second is to feed theory with useful description. One reason why we know little about broad sets of abilities is that adaptive logic is yet to be developed for some abilities. However, the lack of such logic this does not imply the presence or absence of a functional link. A complementary approach is to explore, describe, and follow up associations between adversity and abilities to aid theory development. Therefore, we return to the map of cognitive abilities that might be shaped by adversity and ask “what territory needs exploration and which areas may need re-mapping?”

To carefully examine and interpret data in a principled exploration, it is helpful to develop inferential criteria. For example, rather than using adaptive-logic to predict which abilities are enhanced or reduced, we can ask what criteria are needed for evaluating and interpreting different data patterns? In addition, research typically focuses on reduced versus enhanced test performance, but some abilities might remain ‘intact’ (unaffected) by exposure to adversity (Frankenhuis, Young, et al., 2020). We know little about the intact abilities of people exposed to adversity. We also know little about the drivers of reduced performance on broad and generic measures of ability and achievement. For example, deficit approaches have collapsed many abilities into composites and find that adversity exposure reduces performance. However, one possibility is that a smaller set of specific abilities are driving effects. In total, there is still much to learn about the map of adversity shape cognitive abilities. Principled exploration can complement confirmatory research in drawing this map, especially in the early stages of a new field.

# The Current Study

We conduct a principled exploration of how adversity relates to performance on a widely-used cognitive achievement battery using longitudinal, prospective data from the Study of Early Childcare and Youth Development (SECCYD). Drawing on the general insights of adaptation-based research, we employ a within-person performance design to explore performance across 10 abilities. This design allows us to assess how exposure to each measure of adversity are associated with relative performance differences across many abilities (see Figure 1). In other words, we can compare specific abilities (e.g., short-term memory performance) to overall performance (within-person average performance on all tests) to get a clear picture of how enhanced and reduced performance manifest in parallel within an individual.

We focus on adversity measures that tap two constructs: environmental harshness and unpredictability. We focus on these constructs because they feature often in adaptation-based research on cognitive abilities (Ellis et al., 2017; Ellis et al., 2020; Fields et al., 2021; Frankenhuis, Young, et al., 2020; Mittal et al., 2015; Young et al., 2018; Young et al., 2022). We use both classic and unexplored measures that capture both. Classic measures include family transitions (paternal, residential, and occupational changes) and income-to-needs (mean and variability). Unexplored measures include neighborhood disadvantage (mean and variability). We leverage data from the 1990 Census about the broader ecological context, which has been used to measure the neighborhood context in the SECCYD previously (Bleil, Appelhans, et al., 2021; Bleil, Spieker, et al., 2021).

We outline two sets of criteria for evaluating results. First, our expectations change according to the conceptual framework. For example, from a traditional deficit perspective, we should expect negative overall effects of adversity. Performance on subtests should closely match the overall effect. In contrast, from an adaptation-based perspective, we expect an overall negative effect but performance on some subtests is either less reduced, intact, or even enhanced.

Our second set of criteria are statistical. Our modeling strategy allows us to quantify performance as a function of adversity in two ways. First, we can test whether the effect of adversity on each subtest is different from zero using a simple slopes test. A positive and negative effect suggests enhanced and reduced performance, respectively. Second, we compare subset performance (simple slope) against overall performance (main effect of adversity across all tests), which is measured by the interaction between subtest category and adversity. This interaction term indicates whether performance is significantly more negative, less negative, or even positive compared to overall performance. For both types of effects, we can determine if they are practically equivalent to either zero (simple effect) or overall performance (main effect). Subtest performance is intact when the effect of adversity on a subtest is practically equivalent to zero. Using these criteria, we position ourselves to identify the drivers of reduced overall cognitive performance, map out sets of ‘intact’ cognitive abilities, and discover (possible) enhancements.

|  |
| --- |
| Figure 1 |

# Method

## Participants

Families were initially recruited for the NICHD SECCYD in 1991. A total of 1364 families met all the prescreening criteria, namely that mothers: (a) were age 18 or older, (b) did not plan to move, (c) had a newborn without any known disabilities (and could leave the hospital within one week), (d) had no history of substance abuse, (e) could speak English, and (f) lived within 1 hour driving distance from the research lab and were in a relatively safe neighborhood. More information about recruitment and selection procedures is available from the study [Network (2005); see https://www.icpsr.umich.edu/web/ICPSR/series/00233]. The current analyses included participants with non-missing data on most predictors and outcome variables through age 15 (N = 1156).

## Measures

### **Cognitive Ability Test Battery**

We used the Woodcock-Johnson (WJ) Cognitive and Achievement standardized test battery to examine performance across 10 subtests (Woodcock, 1990; Woodcock et al., 1990). The SECCYD administered the WJ five times over the 54 month, 1st grade, 3rd grade, 5th grade, and 15-year assessments.

There are two WJ test batteries, the cognitive and achievement tests. The WJ cognitive test includes the Memory for Names, Memory for Sentences, Verbal Analogies, Incomplete Words, and Picture Vocabulary subtests (see Table 1 and below for descriptions). The WJ achievement battery includes Letter-Word Identification, Passage Completion, Calculations, Applied Problems, and Word Attack subtests (see below for descriptions).

For all tests, we analyzed standard scores, which are equivalent to IQ scores (e.g., M = 100, SD = 15). Using standard scores for subtests puts all tests on the same scale to facilitate comparison (see Figure 2). For each subtest, we averaged standard scores over time to create one score per subtest per participant. However, the specific set of subtests administered at each assessment varied (see Figure 2). For example, the Verbal Analogies test was measured at grade three and age 15 whereas Passage completion was measured at grades 3, 5, and age 15 (see Table 1). Thus, to create overall scores for each subtest, we averaged over all timepoints available for each subtest.

|  |
| --- |
| Figure 2 |

**Picture Vocabulary.** This subtest measures verbal comprehension and crystallized knowledge. The test contains 58 items requiring participants to view and name familiar and unfamiliar objects. The test was administered five times at 54 months, grades 1, 3, 5, and at 15 years. Higher scores indicate more verbal comprehension and more crystallized knowledge.

**Verbal Analogies.** This subtest measures the ability to reason about analogies between relatively simple words. Although the words remain simple, relations between words increase in complexity of over the test items. The test contains 35 items and was assessed twice at grades 3 and 5. Higher scores indicate more reasoning and more verbal/crystalized knowledge.

**Passage Comprehension.** This subtest test measures the ability to read a short passage and name an appropriate key word that is missing. The test contains 43 items and was administered three times at grades 3, 5, and at age 15. Higher scores indicate more vocabulary, comprehension, and reading skill.

**Applied Problems.** This subtest contains a set of practical math problems. Participants must read and identify a strategy for solving the problem and execute simple arithmetic calculations. The test contains 60 items and was administered at the 54-month, 1st, 3rd and 5th grade, and 15-year assessments. Higher scores indicate more practical math and problem-solving skill.

**Calculations.** This subtest requires participants to solve traditional math problems containing addition, subtraction, multiplication, division, and different combinations of each. The test also includes some geometry and trigonometry problems. Some items require logarithmic operations and calculus. The test contains 58 items and was administered at the 3rd and 5th grade assessments. Higher scores indicate more mathematical/quantitative skill.

**Memory for Names.** This subtest is an auditory-visual association test. It requires participants to learn a set of ‘space creatures’ and their names. After learning a set of creature-name pairs, participants must identify which of a set of nine creatures were just introduced and those previously in past sets. The test difficulty is controlled by (decreasing) increasing the create-name pairs presented in each set. The test contains 72 items and was administered at the 1st and 3rd grade assessments. Higher scores indicate more visual-auditory association and long-term memory skill.

**Incomplete Words.** This subtest measures the ability to listen to words containing missing phonemes and complete the word. The test contains 40 items and was administered at the 54 month and 1st grade assessments. Higher scores indicate more auditory processing skill.

**Memory for Sentences.** This subtest measures the ability to listen to and remember words, phrases, and sentences. The words, phrases, and sentences are played on an audio tape and participants must recall as many as possible. The test contains 32 items and was administered at the 54-month, 1st grade, and 3rd grade assessments. Higher scores indicate more short-term memory skill.

**Letter-word Identification.** This subtest measures reading and pronunciation ability. Participants must initially read letters and then words, which gradually increase in difficulty. The test contains 57 items and was administered at the 54-month, 1st, 3rd, and 5th grade assessments. Higher scores indicate more verbal knowledge.

**Word Attack.** This subtest measures the ability to pronounce unfamiliar words. Participants must read aloud phonetically logical but nonsense or infrequent words. It contains 30 items and was administered at the 1st and 3rd grade assessments. Higher scores indicate more auditory processing and linguistic structural analysis knowledge and skill.

**Table** **:** Bivariate correlations and descriptive statistics for WJ subtests.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Passage Comprehension | - |  |  |  |  |  |  |  |  |  |
| 2. Picture Vocab | 0.66\*\* | - |  |  |  |  |  |  |  |  |
| 3. Calculations | 0.75\*\* | 0.75\*\* | - |  |  |  |  |  |  |  |
| 4. Verbal Analogies | 0.66\*\* | 0.73\*\* | 0.76\*\* | - |  |  |  |  |  |  |
| 5. Letter-Word Pronunciation | 0.61\*\* | 0.56\*\* | 0.66\*\* | 0.64\*\* | - |  |  |  |  |  |
| 6. Short-Term Memory | 0.49\*\* | 0.43\*\* | 0.48\*\* | 0.51\*\* | 0.50\*\* | - |  |  |  |  |
| 7. Applied Problems | 0.45\*\* | 0.42\*\* | 0.47\*\* | 0.43\*\* | 0.35\*\* | 0.30\*\* | - |  |  |  |
| 8. Auditory Processing | 0.64\*\* | 0.63\*\* | 0.78\*\* | 0.71\*\* | 0.57\*\* | 0.47\*\* | 0.50\*\* | - |  |  |
| 9. Unfamilar Words | 0.48\*\* | 0.55\*\* | 0.67\*\* | 0.59\*\* | 0.47\*\* | 0.39\*\* | 0.43\*\* | 0.84\*\* | - |  |
| 10. Auditory-Visual Associations | 0.50\*\* | 0.58\*\* | 0.65\*\* | 0.76\*\* | 0.46\*\* | 0.43\*\* | 0.39\*\* | 0.65\*\* | 0.60\*\* | - |
| N | 1156 | 1064 | 1080 | 1155 | 1142 | 1092 | 1103 | 1154 | 1103 | 1075 |
| Mean | 102.64 | 111.08 | 108.63 | 107.91 | 97.04 | 96.21 | 105.39 | 106.94 | 106.84 | 112.10 |
| SD | 13.65 | 15.65 | 12.85 | 13.62 | 14.68 | 10.74 | 13.32 | 13.14 | 14.30 | 16.72 |
| Min | 31.50 | 50.00 | 42.00 | 33.50 | 45.00 | 39.00 | 46.00 | 30.00 | 54.00 | 12.00 |
| Median | 102.55 | 110.25 | 108.67 | 108.67 | 96.33 | 96.50 | 105.00 | 107.00 | 106.00 | 113.50 |
| Max | 147.00 | 156.00 | 146.00 | 149.00 | 144.00 | 128.00 | 162.00 | 150.00 | 152.00 | 178.50 |

### Indicators of Harshness

We measured environmental harshness in two ways. First, following previous studies using data from the SECCYD, we used income-to-needs ratio scores from 1, 6, 15, 24, 36, and 54-month assessments (Belsky et al., 2012; Hartman et al., 2018; Li et al., 2018; Sung et al., 2016; Zhang et al., 2022). We calculated a simple average of all income-to-needs scores across assessments (M, SD).

Second, we leveraged data from the 1990 Census about the broader economic and ecological context in a similar way to previous analyses of neighborhood-level economic conditions in the SECCYD (Bleil, Appelhans, et al., 2021; Bleil, Spieker, et al., 2021). Specifically, addresses were tracked for each participant over time. Each family address start and stop dates were recorded, geocoded, and linked to the 1990 decennial Census blocks. These blocks are the smallest Census-tracked geographical unit. For each Census block, sociodemographic data were extracted to measure neighborhood-level economic conditions for each participant. We extracted 5 variables: 1) percent of people living under the poverty line, 2) median household income, 3) Gini coefficients of income inequality based on income bracket frequency data, 4) percent of unemployed individuals over 16 in the workforce, and the percent of occupied houses that were being rented. These neighborhood variables first were standardized and averaged within each unique Census block a participant lived in and then averaged over time (up until the 54-month assessment). Thus, if a participant lived in two homes between birth and the 54-month assessment, neighborhood-level variables would be standardized and averaged within the first and second Census block and then averaged between them. These scores served as neighborhood-level socioeconomic harshness where higher scores indicate higher rates of poverty, income-inequality, unemployment, and lower education.

### **Indicators of Unpredictability**

Environmental unpredictability is notoriously hard to define and measure (Young et al., 2020). Nonetheless, studies leveraging data from the SECCYD have used two approaches. The first is track and count family transitions, including changes in paternal figures, parental job transitions, and residential changes (Belsky et al., 2012; Hartman et al., 2018). The second approach is to quantify variability in repeated measures of harshness indicators (e.g., computing measures of variance across time). For example, Li and colleagues (2018) fit a linear model to each participants’ income-to-needs scores over time. Then they computed the residual variance around participant-level linear trends in income-to-needs to create score tapping income variability. In the current study, we compute unpredictability scores using both approaches and extend the Li and colleagues (2018) approach to the neighborhood-level Census block data.

To calculate a family transitions, we computed the number of paternal figure, mother and father (figure) job, and residential changes across 17 assessments from 1 to 54 months (Belsky et al., 2012; Hartman et al., 2018). After computing scores across time, we standardized each variable and averaged them together to compute an overall family transitions variable.

We calculated both familial and neighborhood-level economic variability. For, income-to-needs scores, we computed the simple standard deviation of all income-to-needs scores for each participant from the 1, 6, 15, 24, 36, and 54-month assessments. For neighborhood-level economic variability, we computed the standard deviation of neighborhood socioeconomic harshness scores (see Indicators of Harshness). If participants had only lived in one Census block from 1 to 54 months, their neighborhood socioeconomic harshness variability score was 0.

### **Control Variables**

We used a standard set of three control variables typically used in analyses of SECCYD data: 1) maternal education, 2) sex assigned at birth (1 = female), and 3) the race/ethnicity of each child coded as White/non-Hispanic = 0, otherwise = 1.

**Table** **:** Bivariate correlations and descriptive statistics for adversity variables.

| Variable | 1 | 2 | 3 | 4 | 5 |
| --- | --- | --- | --- | --- | --- |
| 1. Family Transitions | - |  |  |  |  |
| 2. Income-to-Needs (Mean) | -0.29\*\* | - |  |  |  |
| 3. Income-to-Needs (SD) | -0.11\*\* | 0.70\*\* | - |  |  |
| 4. Neighborhood (Mean) | 0.24\*\* | -0.46\*\* | -0.25\*\* | - |  |
| 5. Neighborhood (SD) | 0.45\*\* | -0.17\*\* | -0.07\* | 0.29\*\* | - |
| N | 1351 | 1355 | 1299 | 1295 | 1295 |
| Mean | -0.01 | 3.39 | 1.09 | -0.00 | 0.00 |
| SD | 0.70 | 2.69 | 1.14 | 0.75 | 0.80 |
| Min | -0.82 | 0.13 | 0.00 | -2.22 | -0.69 |
| Median | -0.19 | 2.72 | 0.81 | -0.08 | -0.18 |
| Max | 3.86 | 23.79 | 17.78 | 3.35 | 3.15 |

# Results

## Preregistration, Statistical Power, and Computational Reproducibility

We preregistered this study using a template for secondary data analysis (Akker et al., 2021). The preregistration document and its entire version history was tracked on GitHub (see <https://tinyurl.com/seccyd-wj-prereg> for the document and <https://tinyurl.com/seccyd-wj-prereg-history> for revision history).

We also conducted a power analysis as part of our prergistration (see <https://tinyurl.com/seccyd-wj-power> for write up and see <https://tinyurl.com/seccyd-wj-power-code> for code). In short, we used a simulation approach to conduct power analyses. Although we simulated adversity scores, we leveraged actual WJ test scores from the data. Simulations showed 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.

All relevant files (data processing, analysis code, manuscript etc.) for this project are tracked on GitHub (see <https://tinyurl.com/seccyd-wj>), including data needed to reproduce all results (see <https://tinyurl.com/seccyd-wj-data>). Raw data (data provided by SECCYD) is only available via ICPSR ([https://www.icpsr.umich.edu](https://www.icpsr.umich.edu/rpxlogin)). However, documentation for the study is free to download (see <https://www.icpsr.umich.edu/web/ICPSR/studies/21940>), which contains lists of raw datasets and variables. For those who have access to raw SECCYD data, we provide a table of raw datasets and variables used in this project (see <https://tinyurl.com/seccyd-wj-data>).

We used R , Rstudio, and Quarto to process, analyze, and report results (Allaire, 2022; Posit team, 2023; R Core Team, 2023). For reading raw SECCYD data, used the haven and readxl R packages (Wickham et al., 2023; Wickham & Bryan, 2023). For data processing, visualizations, and table creation, we used the tidyverse, sjlabelled, ggdist, and patchwork R packages (Gohel & Skintzos, 2023; Kay, 2023; Lüdecke, 2022; Pedersen, 2022; Wickham et al., 2019). For analyses, including mixed models, simple slopes, and equivalence tests, we used lme4, faux, ggeffects, marginaleffects, and parameters R packages (Arel-Bundock, 2023; Bates et al., 2015; DeBruine, 2023; Lüdecke, 2018; Lüdecke et al., 2020).

## Data Analysis Strategy and Inferential Criteria

We used a mixed effecrts modeling approach to analyze how adversity relates to WJ performance. For our primary analyses, we ran one model per adversity variable. Each model contained sex assigned at birth, race/ethnicity, and maternal education as covariates. Adversity and covariates were standardiazed or recoded to center variables at 0.

To analyze and compare WJ subtest performance with overall WJ performance, we restructured the data so that each participant was represented by 10 rows, one for each WJ subtest score. Then, we created a sum coded contrast variable for WJ subtests with 10 levels (one for each subtest). This type of contrast sets the model interecpt to the grand mean (e.g., the mean of all subtest scores). To analyze the effects on adverstiy on test performance, we entered adversity as a main effect and the interaction between adversity and the contrast-coded subtest variable. A model with this structure will contain a main effect for each covariate and one for adversity and an interaction term for each subtest (i.e., 10 interaction terms). The main effect of adversity reflects the association between adversity and overall WJ performance (e.g., within-person average of all subtests; see Figure 1). The interaction terms reflect the association between adversity and subtest performance. However, interactions reflect the difference between the overall performance and subtest performance (see Figure 1). Thus, the interaction measures whether specific WJ subtest performance differs from overall test performance as a function of adversity exposure.

Using this modeling strategy, we are 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; and 3) the interaction effect between an adversity measure and subtest. We do not have specific point or range predictions for the effect size types above. However, we view standardized regression coefficients (i.e., ) .10 (or higher) and -.10 (or lower) meaningful. For main effects, effects outside this range indicate that overall performance is meaningfully positve or negative across levels of adversity. For interactions, effect sizes outside these bounds indicate that a subtest score is meaningfully more negative or more positive compared with the overall score. For simple effects, effects outside these bounds indicate that the effect of adversity on a specific subtest is meaningfully different from zero. We are also interested in null effects. Specifically, we use equivalence testing to determine if a given effect is practically equivalent to a Range of Practical Significance (ROPE). We chose a ROPE falling between = -.10 and = .10 (Kruschke, 2018; Lakens et al., 2018). Although we report standardized coefficients, we converted our ROPE to the WJ standard score scale by multiplying the standard deviation of standard WJ scores (*SD* = 15) by .1. This means our ROPE was -1.5 to 1.5 for unstandardized coefficients.

To guide interpretation, we apply a set of inferential criteria for categorizing data patterns. We are interested in three data patterns: 1) enhanced performance, 2) reduced performance, and 3) inact performance. We infer ‘enhanced performance’ when main and simple effects are meaningfully positive and statistically different from zero. We infer ‘reduced performance’ when main and simple effects are meaningfully negative and statistically different from zero. We use equivalence tests to interpret patterns of ‘intact performance’. We infer intact performance when a main or simple effect (and its confidence bounds) is practically equivalent to zero (e.g., falls between = -.10 and = .10 ).

We use the same criteria for interaction terms with one difference. Because interaction terms test the difference between main and simple effects, they quantify relative performance patterns. For ‘enhanced relative performance’, interactions terms must be meaningfully positive and significant. For ‘reduced relative performance’, an interaction term must be meaningfully negative and significant. Interaction terms that are practically equivalent to zero reflect subest performance that closely resembles overall perforamnce. However, inferring ‘enhanced’, ‘reduced’, or ‘intact’ relative performance depends on the size and direction of the main effect. We are particularly interested in cases where a main effect is negative and interaction terms is positive. This may reflect ‘enhanced relative performance’ (e.g., meangingful and significant postive interactions), or ‘less reduced’ performance on a particular subtest in the context of an overall reduced pattern of performance.

## Primary Analyses

Our primary analyses examined how indicators of harshness and unpredictability were associated with WJ overall and subtest performance. We ran one mixed model per indicator for a total of five primary analyses (2 for harshness and 3 for unpredictability). All analyses controlled for the main effects of maternal education, race/ethinicty, and sex assigned at birth. Across all models, there were main effects for both maternal education and race/ethnicity. Lower maternal education and having a non-White racial/ethnic background was associated with lower WJ overall performance. No model contained effects for sex assigned at birth. Below we describe the effects of our primary analysis predictors (see Supplement Table 1).

### Indicators of Harshness

**Income-to-Needs (Mean)**. Our mixed-model analyzed the effect of income-to-needs on overall compared with subtest WJ performance. There was a main effect of income-to-needs such that a higher poverty (lower income-to-needs) was associated with lower overall WJ performance. Equivalence tests show that this overall main effect was outside the ROPE (see Figure 3a).

Interaction effects between income-to-needs and subtests revealed a more nuanced landscape of associations. Performance on Passage Completion, Calculations, Verbal Analogies, Letter-Word, Short-Term Memory, and Unfamiliar Words subtests did not differ from the overall main effect (see Figure 3a and 3b). However, performance on the Picture Vocabulary subtest was significantly more negative and practically unequivalent to the overall main effect (see Figure 3a and 3b). Interestingly, performance on the Auditory Processing, Unfamiliar Words, and Auditory-Visual Associaitons subtest were significantly more postive than the overall main effect (see Figure 3a and 3b). However, equivalence tests suggest that performance on the Unfamiliar words test was inside the ROPE, and thus practically equivalent to the main effect.

Our simple effects analysis tested whether the association between income-to-needs was statistically different from zero and whether it was practically equivalent to the ROPE (see Figure 3c). Analyses revlealed that the association between adversity and each of the subtests where significantly negative and outside the ROPE for except for the Auditory Processing, Unfamiliar Words, and Auditory-Visual Associaitons subtests (see Figure 3c). For these tests, all effects were not statistically different from zero and practically equivalent to the ROPE (see FIgure 3c).

Based on our inferential criteria, the main effect of income-to-needs was consistent with reduced overall performance. For the Picture Vocabulary subtest, performance was significantly more reduced than the overall pattern and outside the ROPE. However, three tests showed relative enhancemet to the overall pattern: Auditory Processing, Unfamiliar Words, and Auditory-Visual Associaitons subtests. Yet, only the Auditory Processing and Auditory Visual Associations subtests were outside the ROPE. However, no subtest performance pattern was consistent with pure enhancement. Instead, simple effect analyses revealed that Auditory Processing, Unfamiliar Words, and Auditory-Visual Associaitons performance was inside the ROPE, consistent with intact performance among families with low income.

**Neighborhood Poverty (Mean)**. There was also a main effect of neighborhood poverty such that a living in a high poverty neighborhood was associated with lower overall WJ performance (see Figure 3d). Equivalence tests show that this overall main effect was outside the ROPE.

Interaction effects between neighborhood poverty and subtest were again varied. Performance on Passage Completion, Calculations, Letter-Word, and Short-Term Memory subtests did not statistically differ from the overall main effect (see Figure 3d adn 3e). However, performance on the Picture Vocabulary, Verbal Analogies, and Applied Problems subtests were significantly more negative than the main effect (see Figure 3d and 3e). However, equivalence tests showed that only the Verbal Analogies subtest performance was outside the ROPE (see Figure 3e). Similar to the income-to-needs analysis, performance on the Auditory Processing and Auditory-Visual Associaitons subtests were significantly more postive than the overall main effect. Equivalence tests revealed that performance on both were also outside the ROPE (see Figure 3e).

Our simple effects analysis showed that higher neighborhood poverty was associated with statistically negative and outside the ROPE for all subtests except for the Auditory Processing nd Auditory-Visual Associaitons subtests. Again, for these two subtests, performance was not statistically different from zero and inside the ROPE.

According to our inferential criteria, our analysis suggest that the main effecd of neighbohood poverty consistent with reduced overall performance. For the Verbal Analogies subtest, performance was significantly more reduced and outsdide the ROPE. However, performance on the Auditory Processing and Auditory-Visual Associaitons subtests were consistent with relative enhancement. However, similar to the income-to-needs results, no tests showed pure enhancement. Simple effect analyses revealed mostly reduced performance. For Auditory Processing and Auditory-Visual Associaitons performance showing relative enhancement, simple effects suggest that they are practically equivalent to the ROPE, consistent with an intact pattern of performance among those living in high poverty neighborhoods.

|  |
| --- |
| Figure 3 |

### Indicators of Unpredictability

**Family Transitions**. Our analysis of family transitions revealed no main effect on overall WJ performance. The main effect also fell inside the ROPE range suggesting that overall performance was not associated with exposure to more family transitions (see Figure 4a).

Three interaction terms were statistically different from the main effect: Calculations (more negative), Auditory Processing (more positive), and Audio-Visual Associations (more postive). However, only the effect of family transitions on the Calculations subtest was outside the ROPE (see Figure 4b).

Simple effects revealed there were no associations between family transitions and subtest performance. Only Calculations subtest performance was statistically different from zero and outside the ROPE. Applied Problems subtest performance was outside the ROPE but not statistcally different from zero.

Our inerential criteria suggest that WJ performance is mostly intact among those exposed to family transitions. This was true both for overall WJ performance and most subtets, except for the Calculations subtet, which was consitent with a reduced pattern of performance.

**Income-to-Needs Variability.** Models unpacking the effect of income-to-needs variablity on WJ overall and subtest performance raised questions about the validity of income-to-needs variability scores as indicators of unpredictability. Specifically, the direction of all effects were opposite to analyses using income-to-needs average scores. Simple effects suggest that subtest performance that were reduced by income-to-needs mean scores were enhanced in analyses using income-to-needs variablity. We belieive such effects are driven by the fact that income-to-needs mean and variability scores are strongly related (see Table 2), which has been reported before elsewhere (Li et al., 2018). That is, more poverty was related to less income-to-needs variablity. Put differently, richer families were more likely to experience income fluctuations.

To resolve this issue, we conducted a set of secondary analyses that use different methods for computing variability. We report analyses using income-to-needs varaibility scores and income-to-needs average percent change scores (see Secondary Analyses).

**Neighborhood Poverty Variability**. In contrast to income-to-needs variablity, neighbhorhood poverty variability was related to average neighbhorhood poverty in the expected direciton. That is, families living in poor neighborhoods (more harsh) were more likely to experience variablity in neighborhood poverty (unpredictability) from one to 54 months. Additionally, the association between average and variablity scores were moderate rather than strong (see Table 2).

There was no main effect of neighborhood poverty variablity on overall WJ scores (see FIgure 4). There was only one interaction with subtest performance. Audio-Visual Associations subtest performance was significantly more positive than the main effect, however, this effect was inside the ROPE. In addition, simple effects showed no subtest performance was different from zero and all effects were inside the ROPE.

|  |
| --- |
| Figure 4 |

# 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>

Allaire, J. (2022). *Quarto: R interface to ’quarto’ markdown publishing system*. <https://CRAN.R-project.org/package=quarto>

Arel-Bundock, V. (2023). *Marginaleffects: Predictions, comparisons, slopes, marginal means, and hypothesis tests*. <https://CRAN.R-project.org/package=marginaleffects>

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using Lme4. *Journal of Statistical Software*, *67*(1). <https://doi.org/gcrnkw>

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., Appelhans, B. M., Thomas, A. S., Gregorich, S. E., Marquez, N., Roisman, G. I., Booth-LaForce, C., & Crowder, K. (2021). Early life predictors of positive change during the coronavirus disease pandemic. *BMC Psychology*, *9*(1), 83. <https://doi.org/gk5dzv>

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.

DeBruine, L. (2023). *Faux: Simulation for factorial designs*. <https://doi.org/10.5281/zenodo.2669586>

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2017). Moving Beyond Correlations in Assessing the Consequences of Poverty. In S. T. Fiske (Ed.), *Annual Review of Psychology, Vol 68* (Vol. 68, pp. 413–434). Annual Reviews. <https://doi.org/10.1146/annurev-psych-010416-044224>

Duquennois, C. (2022). Fictional Money, Real Costs: Impacts of Financial Salience on Disadvantaged Students. *American Economic Review*. <https://doi.org/10.1257/aer.20201661>

Ellis, B. J., Abrams, L. S., Masten, A. S., Sternberg, R. J., Tottenham, N., & Frankenhuis, W. E. (2020). Hidden talents in harsh environments. *Development and Psychopathology*, 1–19. <https://doi.org/gjqwbs>

Ellis, B. J., Bianchi, J., Griskevicius, V., & Frankenhuis, W. E. (2017). Beyond risk and protective factors: An adaptation-based approach to resilience. *Perspectives on Psychological Science*, *12*(4), 561–587. <https://doi.org/gdtj9h>

Farah, M. J., Shera, D. M., Savage, J. H., Betancourt, L., Giannetta, J. M., Brodsky, N. L., Malmud, E. K., & Hurt, H. (2006). Childhood poverty: Specific associations with neurocognitive development. *Brain Research*, *1110*(1), 166–174. <https://doi.org/fv2dn8>

Fields, A., Bloom, P. A., VanTieghem, M., Harmon, C., Choy, T., Camacho, N. L., Gibson, L., Umbach, R., Heleniak, C., & Tottenham, N. (2021). Adaptation in the face of adversity: Decrements and enhancements in children’s cognitive control behavior following early caregiving instability. *Developmental Science*, *24*(6), e13133. <https://doi.org/gj7gh7>

Flournoy, J. C., Vijayakumar, N., Cheng, T. W., Cosme, D., Flannery, J. E., & Pfeifer, J. H. (2020). Improving practices and inferences in developmental cognitive neuroscience. *Developmental Cognitive Neuroscience*, *45*, 100807. <https://doi.org/gnbxdn>

Frankenhuis, W. E., de Vries, S. A., Bianchi, J., & Ellis, B. J. (2020). Hidden talents in harsh conditions? A preregistered study of memory and reasoning about social dominance. *Developmental Science*, *23*(4), e12835. <https://doi.org/ggb8qd>

Frankenhuis, W. E., & de Weerth, C. (2013). Does Early-Life Exposure to Stress Shape or Impair Cognition? *Current Directions in Psychological Science*, *22*(5), 407–412. <https://doi.org/f5cxhb>

Frankenhuis, W. E., & Nettle, D. (2020). The Strengths of People in Poverty. *Current Directions in Psychological Science*, *29*(1), 16–21. <https://doi.org/ggf5d6>

Frankenhuis, W. E., Young, E. S., & Ellis, B. J. (2020). The hidden talents approach: Theoretical and methodological challenges. *Trends in Cognitive Sciences*, *24*(7), 569–581. <https://doi.org/10.1016/j.tics.2020.03.007>

Gohel, D., & Skintzos, P. (2023). *Flextable: Functions for tabular reporting*. <https://CRAN.R-project.org/package=flextable>

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659. <https://doi.org/b254c6>

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>

Kay, M. (2023). *Ggdist: Visualizations of distributions and uncertainty*. <https://doi.org/10.5281/zenodo.3879620>

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. <https://doi.org/gfvh58>

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>

Lüdecke, D. (2018). *Ggeffects: Tidy data frames of marginal effects from regression models.* *3*, 772. <https://doi.org/10.21105/joss.00772>

Lüdecke, D. (2022). *Sjlabelled: Labelled data utility functions (version 1.2.0)*. <https://doi.org/10.5281/zenodo.1249215>

Lüdecke, D., Ben-Shachar, M. S., Patil, I., & Makowski, D. (2020). *Extracting, computing and exploring the parameters of statistical models using r.* *5*, 2445. <https://doi.org/10.21105/joss.02445>

McIntosh, R. D. (2017). Exploratory reports: A new article type for cortex. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *96*, A1–A4. <https://doi.org/10.1016/j.cortex.2017.07.014>

McLaughlin, K. A., Weissman, D., & Bitrán, D. (2019). Childhood adversity and neural development: A systematic review. *Annual Review of Developmental Psychology*, *1*(1), 277–312. <https://doi.org/gj59n7>

Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S., & Young, E. S. (2015). Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology*, *109*(4), 604–621. <https://doi.org/10.1037/pspi0000028>

Muskens. (2019). *Hidden obstacles in education for students from low socioeconomic backgrounds:* [PhD thesis, maastricht university]. <https://doi.org/10.26481/dis.20191115mm>

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.

Nweze, T., Nwoke, M. B., Nwufo, J. I., Aniekwu, R. I., & Lange, F. (2021). Working for the future: Parentally deprived Nigerian children have enhanced working memory ability. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *62*(3), 280–288. <https://doi.org/gphn59>

Pedersen, T. L. (2022). *Patchwork: The composer of plots*. <https://CRAN.R-project.org/package=patchwork>

Posit team. (2023). *RStudio: Integrated development environment for r*. <http://www.posit.co/>

R Core Team. (2023). *R: A language and environment for statistical computing*. <https://www.R-project.org/>

Raby, K. L., Roisman, G. I., Fraley, R. C., & Simpson, J. A. (2015). The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 years. *Child Development*, *86*(3), 695–708. <https://doi.org/gjh5cc>

Rifkin-Graboi, A., Goh, S. K.-Y., Chong, H. J., Tsotsi, S., Sim, L. W., Tan, K. H., Chong, Y. S., & Meaney, M. J. (2021). Caregiving adversity during infancy and preschool cognitive function: Adaptations to context? *Journal of Developmental Origins of Health and Disease*, 1–12. <https://doi.org/gphn6c>

Roisman, G. I. (2021). Editorial: A Vision of a Fair and Efficient, Diverse and Inclusive, Cumulative Science of Child Development in the Best and Worst of Times. *Child Development*, *92*(2), 451–465. <https://doi.org/10.1111/cdev.13538>

Rozin, P. (2001). Social Psychology and Science: Some Lessons From Solomon Asch. *Personality and Social Psychology Review*, *5*(1), 2–14. <https://doi.org/bhqn85>

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science*, *16*(4), 744–755. <https://doi.org/ghp4k7>

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>

Ursache, A., & Noble, K. G. (2016). Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology*, *53*(1), 71–82. <https://doi.org/f8jcxn>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). *Welcome to the tidyverse*. *4*, 1686. <https://doi.org/10.21105/joss.01686>

Wickham, H., & Bryan, J. (2023). *Readxl: Read excel files*. <https://CRAN.R-project.org/package=readxl>

Wickham, H., Miller, E., & Smith, D. (2023). *Haven: Import and export ’SPSS’, ’stata’ and ’SAS’ files*. <https://CRAN.R-project.org/package=haven>

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.

Young, E. S., Frankenhuis, W. E., DelPriore, D. J., & Ellis, B. J. (2022). Hidden talents in context: Cognitive performance with abstract versus ecological stimuli among adversity-exposed youth. *Child Development*, *93*(5), 1493–1510. <https://doi.org/10.1111/cdev.13766>

Young, E. S., Frankenhuis, W. E., & Ellis, B. J. (2020). Theory and measurement of environmental unpredictability. *Evolution and Human Behavior*, *41*(6), 550–556. <https://doi.org/10.1016/j.evolhumbehav.2020.08.006>

Young, E. S., Griskevicius, V., Simpson, J. A., Waters, T. E. A., & Mittal, C. (2018). Can an unpredictable childhood environment enhance working memory? Testing the sensitized-specialization hypothesis. *Journal of Personality and Social Psychology*, *114*(6), 891–908. <https://doi.org/10.1037/pspi0000124>

Zhang, X., Schlomer, G. L., Ellis, B. J., & Belsky, J. (2022). Environmental harshness and unpredictability: Do they affect the same parents and children? *Development and Psychopathology*, *34*(2), 667–673. <https://doi.org/gnsdkr>