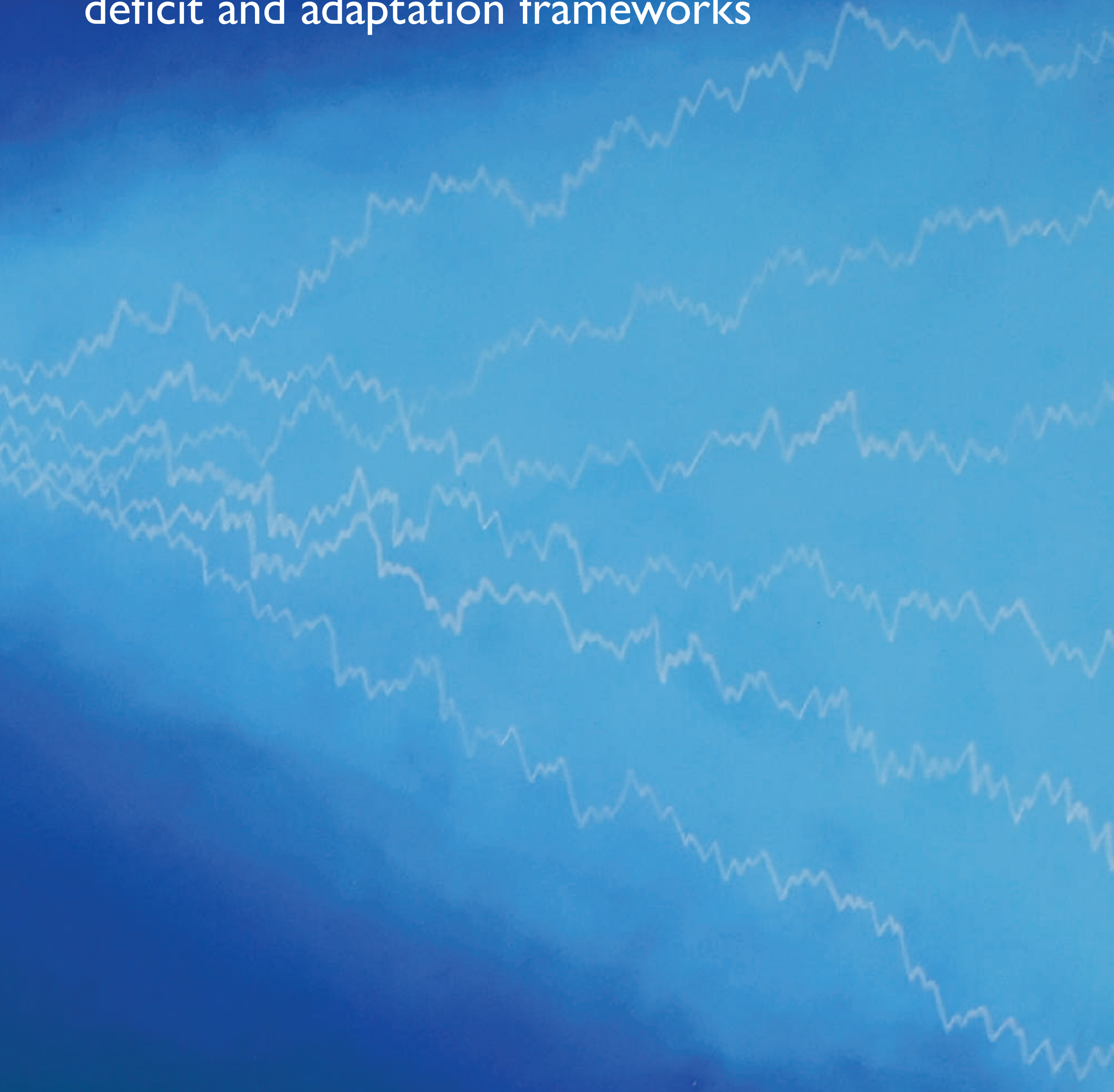
# Executive Functioning in Adverse Environments

## Using cognitive modeling to integrate deficit and adaptation frameworks



Stefan Vermeent

# Executive functioning in adverse environments:

## Using cognitive modeling to integrate deficit and adaptation frameworks

Stefan Vermeent

**ISBN**
978-94-6522-030-7

**Cover design**
Anaïs Thijssen-Vermeent

**Printing**
Ridderprint | www.ridderprint.nl

# Executive functioning in adverse environments

Using cognitive modeling to integrate deficit and adaptation frameworks

**Executieve functies in stressvolle omgevingen: Een integratie van deficit- en adaptatiemodellen door cognitief modelleren**

(Met een samenvatting in het Nederlands)

**Proefschrift**

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. ir. W. Hazeleger,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op

woensdag 16 april 2025 des middags te 4.15 uur

door

**Pieter Cornelis Stefan Vermeent**

geboren op 27 februari 1993
te Gouda

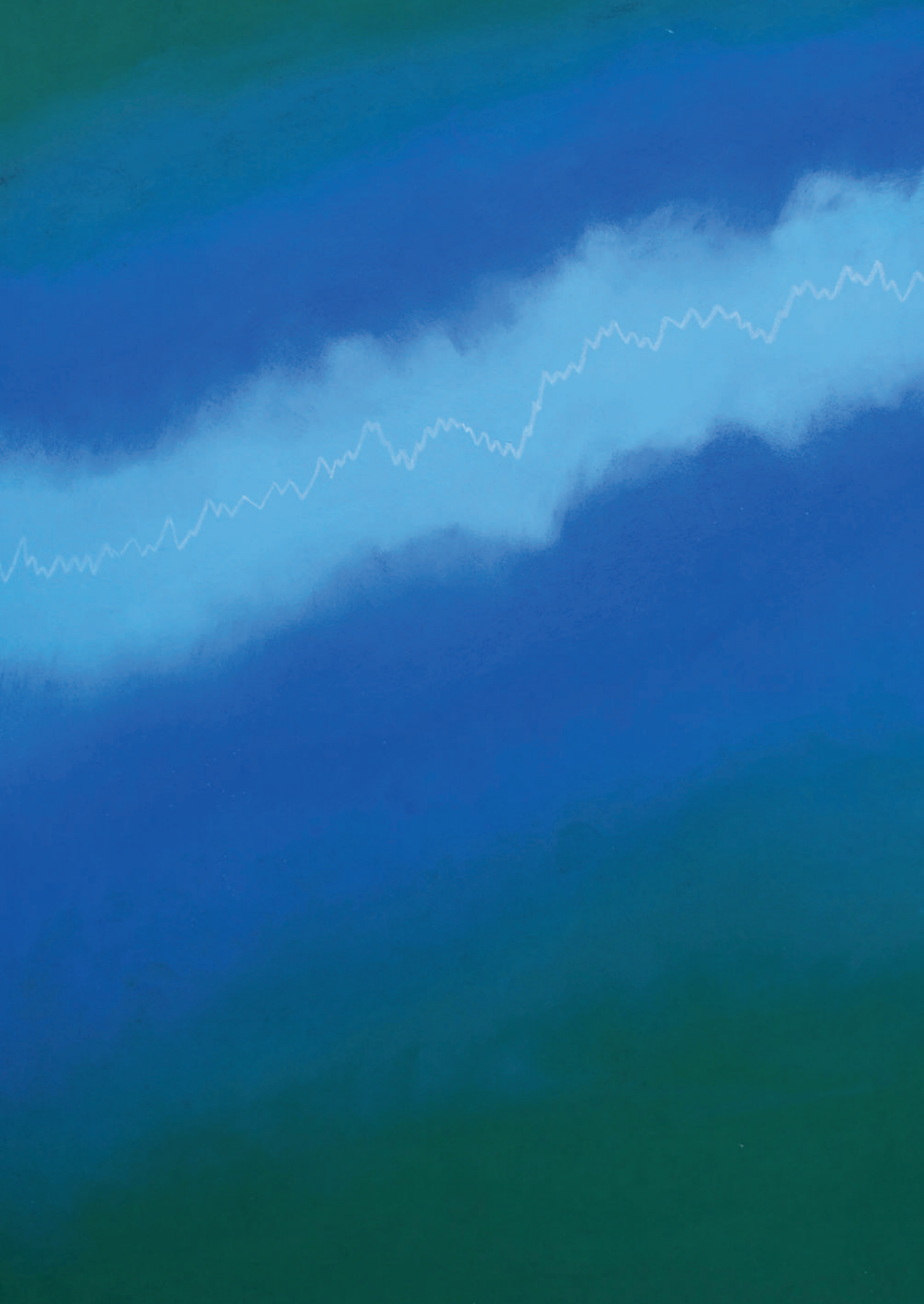
**Promotoren:**
Dr. W.E. Frankenhuis
Prof. dr. J.J.A. Denissen
Prof. dr. J.-L. van Gelder

**Copromotor:**
Dr. E.S. Young

**Beoordelingscommissie**
Prof. dr. N.N.J. Rommelse
Prof. dr. G.W.J.M. Stevens
Prof. dr. S. van der Stigchel
Prof. dr. E.J. Wagenmakers
Dr. A.S. Weigard

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek, de James S. McDonnell Foundation, en ODISSEI.

## Table of Contents



# Chapter 1

General introduction

The field of adversity research is rapidly evolving. Researchers have long been studying the ways in which exposure to adversity impairs cognitive abilities of all sorts, which has led to a proliferation of deficit perspectives. In recent years, however, a growing number of researchers have shifted their attention towards strength-based perspectives. These perspectives highlight the skills, strategies, and knowledge that people may develop in response to adverse experiences. One area in particular that has seen an increase in strength-based thinking is research on executive functioning (EF), which refers to a set of abilities involved in planning, reasoning, and goal-directed behavior.

The next generation of adversity researchers should consider how cognitive deficits and strengths operate alongside each other within the same individual. Studies to date show that people living in adversity will likely exhibit a mix of lower performance in some areas and improved performance in other areas. However, adversity researchers still tend to test deficit and strength-based hypotheses in relative isolation, focusing mainly on one or the other. This has limited a fuller integration of the two perspectives, and therefore prevents a more well-rounded understanding of how adversity shapes cognition. In this dissertation, I will use methodological tools—grounded in mathematical and cognitive psychology—to improve this understanding.

In this chapter, I will first provide an overview of the current state of the field of adversity research, with a special focus on EF. Next, I will discuss methodological issues that limit a fuller understanding of how adversity both lowers and enhances EF. Specifically, I will focus on issues with using standard performance measures as direct measures of EF ability, and explain how cognitive modeling can provide a solution. Finally, I will present an overview of the aims of this dissertation, and the focus of subsequent chapters.

## 1.1 Current state of the field of adversity research

### Exposure to adversity is associated with cognitive deficits

Decades of research have shown that people who experience more adversity—i.e., prolonged exposure to intense stress (for instance, due to violence, deprivation, unpredictability)—tend to score lower on standard cognitive tests (Hackman et al., 2010; Ursache & Noble, 2016a). This lower performance has been documented for a wide variety of cognitive abilities, ranging from executive functioning, social cognition, memory, language, to intelligence (Farah et al., 2006; Sheridan et al., 2022; Sheridan & McLaughlin, 2014). Such findings have led to the proliferation of deficit models, which attribute lower performance in people from adversity to impairments in brain structure and function that undermine social and cognitive abilities (Algarin et al., 2017; Duncan et al., 2010; Farah et al., 2006; Nelson et al., 2020; Nelson & Gabard-Durnam, 2020; Polavarapu & Hasbani, 2017; Rebello et al., 2018; Shonkoff et al., 2012; Ursache & Noble, 2016b). Insights derived from deficit models have informed policy and practice for decades, which have improved the lives of millions of people (Blair & Raver, 2014; Deming, 2009; Duncan et al., 2017; Durlak et al., 2011; Reynolds et al., 2019; Ursache & Noble, 2016a).

### Exposure to adversity is associated with cognitive adaptations

In contrast to deficit frameworks, adaptation frameworks suggest that exposure to adversity could also be associated with intact or even enhanced cognitive abilities. Specifically, people may develop cognitive abilities that help solve unique challenges posed by adverse environments (Ellis et al., 2017, 2022; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013). In adaptation frameworks, the term *enhanced* refers to an ability that has been improved by developmental adaptation, in a way that can be objectively measured (e.g., faster responses, higher accuracy; Frankenhuis, Young, et al., 2020). *Intact* abilities are abilities that are neither enhanced nor impaired by adversity. Cognitive adaptations could lead to intact rather then enhanced abilities, for instance, when performance is also negatively influenced by other deficits (Frankenhuis, Young, et al., 2020; Young et al., 2024).

An important assumption of adaptation frameworks is that specific types of adversity pose their own unique challenges to the individual, and hence require different adaptations (Ellis et al., 2022; Frankenhuis et al., 2016; Frankenhuis & Weerth, 2013). Therefore, testing adaptation hypotheses requires specificity; measures that combine different types of adversity—such as cumulative adversity scores—might not be associated with cognitive enhancements. Contemporary dimensional models of adversity posit that different adversities can be broadly clustered into threat (physical or psychosocial harm), material deprivation (low quantity and quality of material resources), and environmental unpredictability (stochastic variation in adversity, i.e., threat and deprivation, over space and time) (Ellis et al., 2009; McLaughlin et al., 2021; Salhi et al., 2021). Each of these dimensions captures a variety of specific exposures. For instance, exposure to threat may include living with an abusive parent, experiencing high levels of crime in one's neighborhood, or witnessing or participating in fights. Despite this variety, research shows that adversity exposures of the same dimension tend to have similar effects on social and cognitive development, and that their effects are (partially) distinct from effects of other dimensions (Sheridan et al., 2020). Following this work, recent studies have investigated which dimensions of adversity, if any are associated with enhancements in specific EF abilities.

### Developmental adaptations in executive functioning

Adaptation frameworks have sparked a number of studies investigating the development of cognitive abilities in adverse environments (for a review, see Ellis et al., 2022). Several of these have focused on three core components of EF (Karr et al., 2018; Miyake et al., 2000; Zelazo et al., 2013): (1) attention shifting, i.e., efficiently switching between different tasks, (2) working memory updating, i.e., keeping track of changing information in working memory, and (3) inhibition, i.e., ignoring distractions. This line of research hypothesizes that attention shifting and working memory updating are particularly useful abilities in unpredictable and threatening environments. The rationale is that: (a) attention shifting facilitates detecting sudden threats and taking advantage of fleeting opportunities, and (b) working memory updating facilitates tracking changes in the environment. On the other hand, inhibition could actively interfere with detecting and tracking changes in one's environment (Fields et al., 2021; Mittal et al., 2015; Young et al., 2018). Thus, adaptation perspectives predict enhanced performance on attention shifting and working memory updating tasks (in contrast to deficit frameworks), but predict lower performance on inhibition tasks (similar to deficit frameworks).

Several studies have obtained support for adaptation hypotheses, although results are sometimes mixed. Some studies found that exposure to unpredictability (Fields et al., 2021; Mittal et al., 2015; Young et al., 2022) and violence (Young et al., 2022) are positively associated with attention shifting (for counter-examples, see Mezzacappa, 2004; Nweze et al., 2021; Rifkin-Graboi et al., 2021). Two studies found that exposure to unpredictability (Young et al., 2018) and violence (Young et al., 2022) are positively associated with working memory updating. Finally, exposure to different types of adversity, as well as lower socioeconomic status (which is correlated with, but not the same as adversity exposure), have been found to be negatively associated with inhibition (Farah et al., 2006; Mezzacappa, 2004; Mittal et al., 2015; Noble et al., 2005; Rifkin-Graboi et al., 2021). Collectively, these results suggest that exposure to adversity does not uniformly negatively affect EF abilities, but that associations may differ for specific EF abilities.

### Integrating deficit and adaptation frameworks

Deficit and adaptation frameworks are generally considered complementary; within the same person, exposure to adversity could impair some abilities, while enhancing others (Frankenhuis, Young, et al., 2020). However, their integration is still limited. For instance, it is largely unclear which specific abilities may be impaired and which ones may be enhanced by specific types of adversity, and how deficit and adaptation processes may operate alongside each other within the same person. In addition, disentangling deficit and adaptation processes can often be challenging. For example, adaptations in specific abilities may co-occur with general disruptions in brain architecture and neural efficiency due to chronic stress (Shonkoff et al., 2012). As I will argue in the next section, one major methodological challenge limiting a further integration is that both frameworks tend to infer differences in cognitive abilities based on raw performance scores, such as average response times and error rates.

## 1.2 Different reasons for lower performance on EF tasks

Performance on EF tasks is often used as a direct proxy for EF ability, but differences in performance do not necessarily reflect differences in ability. The reason why becomes clear when looking at their dictionary definitions. Performance is defined as the execution of an action (Merriam-Webster, 2024b). In the context of EF tasks, performance may refer to the speed or accuracy of a person's responses to the task. Ability is defined as the natural aptitude for, or acquired proficiency in doing something (Merriam-Webster, 2024a). In the context of EF tasks, this concerns the aptitude for, or acquired proficiency in solving the challenge posed by the task. For performance to equal ability, it is important that the ability is the only factor (or at least a very substantial factor) determining how actions are executed on EF tasks.

However, research in cognitive psychology shows that performance on EF tasks is influenced by cognitive processes other than the specific EF ability that is often of primary interest. Examples of these other processes—discussed in more detail below—are a person's level of response caution, speed of response execution, and general processing speed. This means that two people with the same EF ability level could differ in their performance if, for instance, one of them responds more cautiously than the other (i.e., prioritizing accuracy over speed). These emerging findings, and their implications, have so far mostly been overlooked in adversity research. Given this, long-held assumptions about how adversity affects executive functions could be oversimplified or even misguided.

### A brief case-study

To illustrate the limitations of raw performance scores, imagine a child from a disadvantaged background, let's call her Yara, who is struggling in school. Yara is selected by a screening program designed to proactively identify children who need additional support. The screening includes a brief battery of EF tasks. The results reveal that Yara's response times are below average on nearly all tasks, with particularly low scores on tasks assessing inhibition and working memory. The screening program concludes that Yara has deficits in multiple cognitive abilities. To help her thrive in school, it is recommended that she receive targeted interventions aimed at strengthening her EF, particularly focusing on inhibition and working memory. These interventions might involve cognitive training exercises, tutoring, or behavioral strategies to help her focus and better manage her impulses.

Are these recommendations justified? Perhaps, but, as I will show in the following sections, there are alternative explanations for Yara's lower performance that should be considered. First, Yara may be slower not because of lower EF ability, but because she uses a different strategy than other children, which could affect her performance. For instance, her responses may be more cautious, sacrificing speed to achieve a higher level of accuracy. Second, lower performance across tasks could be driven by a single process common to all tasks, rather than reflecting deficits in specific cognitive abilities. Both issues can give rise to a performance-ability gap, meaning that Yara's raw performance on EF tasks might not accurately reflect her true EF abilities. In the next sections, I will outline the issues with raw performance scores as proxies of cognitive ability, and explain how adversity research can address these issues by building bridges with mathematical and cognitive psychology.

### Limitations of raw performance scores

Research often relies on raw performance scores as a proxy for cognitive abilities. Most often these are response times (i.e., the total time taken to complete a task) and accuracy (i.e., whether the decision made is correct or incorrect). Although there are exceptions, researchers generally focus on one over the other, and these practices can differ between tasks. For instance, performance on inhibition tasks tends to be summarized using the average response time, while performance on working memory updating tasks tends to be summarized using the overall error rate (Bastian et al., 2020). For simplicity, I will mostly focus on response times in this section. However, the same arguments generally also apply to accuracy.

The use of response times is based on the assumption that cognitive operations involve multiple distinct processes, each of which takes time to complete (Donders, 1869). When any of the processes in the chain takes longer to complete, this results in an increased response time. This is also the basic rationale behind commonly used difference scores, where the mean response time of one condition is subtracted from the mean response time of another condition (Donders, 1869). For instance, many EF tasks include a condition with lower processing demands (e.g., trials on the Stroop Task where the color matches the printed word) and a condition requiring the same processing demands *plus* an additional processing demand (e.g., trials on the Stroop trials where the color does not match the printed word). As the conditions are assumed to differ only in terms of the added processing demand, a difference score is thought to isolate the speed of that specific process.

The problem with these approaches is that the use of response times is not based on a formally defined model of how the cognitive system works. Response times are assumed to reflect several cognitive processes, but these processes are treated largely as a black box. This leads to a reverse-inference problem when using response times to infer cognitive ability: just because a lower ability leads to slower response times, does not mean that slower response times reflect lower ability (White & Kitchen, 2022). Common approaches to account for other processes, such as difference scores, have been shown to be insufficient (Miller & Ulrich, 2013). For instance, analyses based on response times fail to account for speed-accuracy trade offs: Some people may take longer to complete a task because they prioritize accuracy over speed, not because they process information more slowly (Bogacz et al., 2010; Van Veen et al., 2008). If adversity exposure is associated with changes in these other processes, the resulting difference in performance could be misattributed to their cognitive abilities.

## 1.3 From performance to cognitive processes: Computational models of cognition

Cognitive modeling offers a fruitful way to bridge the gap between raw performance scores and EF ability (Guest & Martin, 2021; Patzelt et al., 2018). Cognitive models are formalized, mathematical accounts of cognitive processes. They make explicit assumptions about the (unobserved) cognitive processes that give rise to differences in raw performance, and formalize these assumptions in mathematical language. The result is one or more model *parameters* that represent distinct cognitive processes. By applying a cognitive model to empirical performance data, we can generate parameter estimates that best explain key patterns in the data. These parameter estimates can then be used as measures of cognitive processes, and subsequently as predictors or outcomes of interest.

Figure 1.1 shows how a workflow based on cognitive modeling differs from one based on analyzing raw performance. Performance-based workflows infer cognitive abilities either directly from raw performance scores (e.g., mean response time), or by calculating the difference in performance between an experimental and a control condition. Thus, these approaches assume that response times directly reflect the ability of interest. In contrast, a cognitive modeling workflow provides an explicit mathematical account of how performance is shaped by a collection of cognitive processes. Adversity research can use cognitive models to obtain direct measures of the cognitive processes involved in EF tasks, and to investigate if and how they are associated with adversity exposure.

### Cognitive models of decision-making

Most common EF tasks require some kind of binary decision-making: deciding whether an arrow points left or right, classifying a geometric shape either in terms of its color or shape, or deciding whether the currently presented letter is the same as the one presented earlier. These decisions usually have to be made under time pressure, meaning that people have to balance being fast with being accurate. Cognitive models of decision making explain how people make these kinds of decisions, and how they balance demands on speed and accuracy. In cognitive psychology, these models have proven their usefulness for explaining performance on EF tasks relative to raw performance measures (Frischkorn et al., 2019; Hedge et al., 2022; Löffler et al., 2024). For adversity research, they could similarly prove useful in better understanding why people with more exposure to adversity sometimes perform lower, and sometimes perform higher.

### Drift Diffusion Model

One of the most well-validated and successful models of decision-making is the Drift Diffusion Model (DDM; Forstmann et al., 2016; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Wagenmakers, 2009). The DDM accounts for the cognitive processes that give rise to patterns of RTs and error rates (Ratcliff et al., 2015). It models decision-making as an evidence accumulation process, in which people repeatedly sample in-

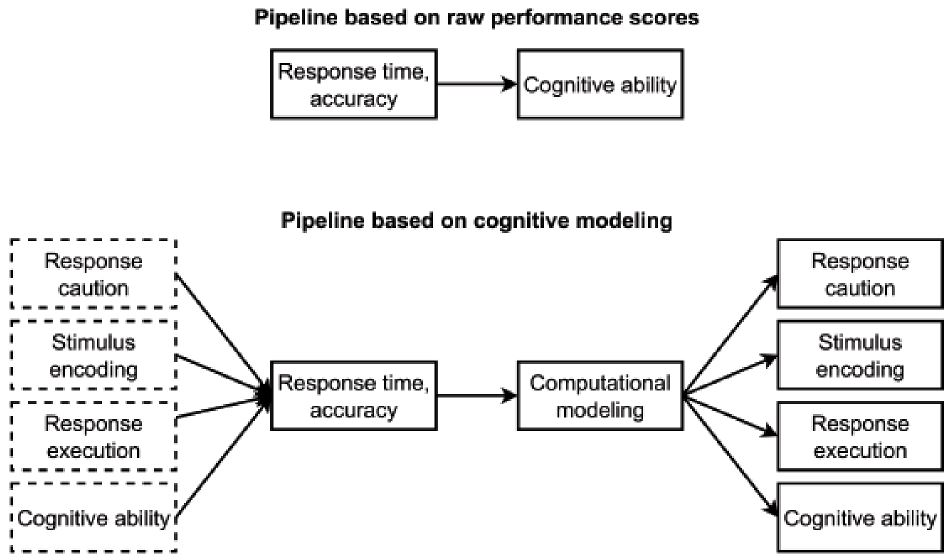



**Figure 1.1.** Workflow based on raw performance (top) versus workflow based on cognitive modeling (bottom).

formation until they have sufficient information to favor one response option over the other (see Figure 1.2). Evidence accumulation is modeled as a random walk process, which drifts towards one of two decision boundaries, usually corresponding to the correct or incorrect response[1]. When the evidence accumulation process reaches one of the two decision boundaries, the response is initiated. The DDM also accounts for the time that it takes to encode stimulus information before evidence accumulation starts, and the time that it takes to execute a response after a decision has been made.

Applying the DDM to trial-level RT and accuracy data yields three parameters[2] that represent distinct cognitive processes. These are (1) the *drift rate*, (2) the *boundary separation*, and (3) the *non-decision time*. Figure 1.3 shows how changes in each DDM parameter shape performance using simulated data. Compared to a baseline for a hypothetical participant, the Figure shows how changes in isolated DDM parameters affect specific aspects of response time distributions.

The *drift rate* is the average rate across trials with which evidence accumulation reaches the correct boundary, and measures the efficiency of evidence accumulation. A decrease in drift rates affects performance in two ways (see Figure 1.3B). First, a lower drift rate increases the spread in the tail of the distribution, coupled with only a small change in the peak of the distribution. Second, it leads to an increased error rate. Thus, people with a lower drift rate respond more slowly and commit more errors. Individual differences in drift rates are generally considered as reflecting differences in cognitive

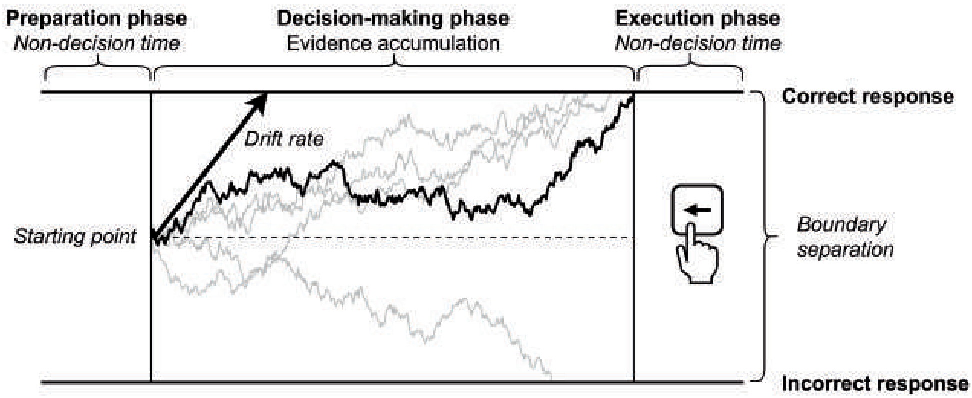



**Figure 1.2.** An overview of the Drift Diffusion Model (DDM). The DDM assumes that people sequentially move through three distinct stages when completing tasks with two forced-response options. First, in a preparation phase, people visually encode stimuli. Second, in the decision phase, people accumulate information favoring one decision over the other (e.g., whether to press the left vs. right key). Each jagged line represents this accumulation on a single trial. Third, in the execution phase, people execute a motor response (e.g., pressing the left vs. right key). The DDM estimates four parameters that represent four distinct cognitive processes (italicized): (1) *Drift rate*: the average rate of evidence accumulation towards the correct decision boundary; i.e., **efficiency of evidence accumulation**; (2) *Non-decision time*: the time spent on processes outside of the decision phase, i.e., **encoding stimuli and executing response**; (3) *Boundary separation*: the distance between decision boundaries; i.e., **response caution**; (4) *Starting point*: the starting point of the decision process; i.e., **response bias**. Figure copied from Vermeent et al. (2024).

ability (Löffler et al., 2024; Schmiedek et al., 2007; Voss et al., 2013). However, as discussed in section 1.4, drift rates also capture general processes (Lerche et al., 2020; Löffler et al., 2024; Weigard et al., 2021).

The *boundary separation* is the width between the two decision boundaries, and measures the level of response caution. An increase in boundary separation affects performance in two ways (see Figure 1.3C). First, a larger boundary separation shifts the peak of the distribution to the right, and also increases the spread of the distribution. Second, it leads to a reduced error rate. Thus, the boundary separation captures the speed-accuracy trade off: a larger boundary separation leads to more accurate yet slower responses.

The *non-decision time* is a combination of the speed of initial stimulus encoding and the speed of response execution. A larger non-decision time shifts the distribution to the right without changing the spread of the distribution and without changing the error rate (see Figure 1.3D). Thus, people with a larger non-decision time are slower without a change in accuracy.

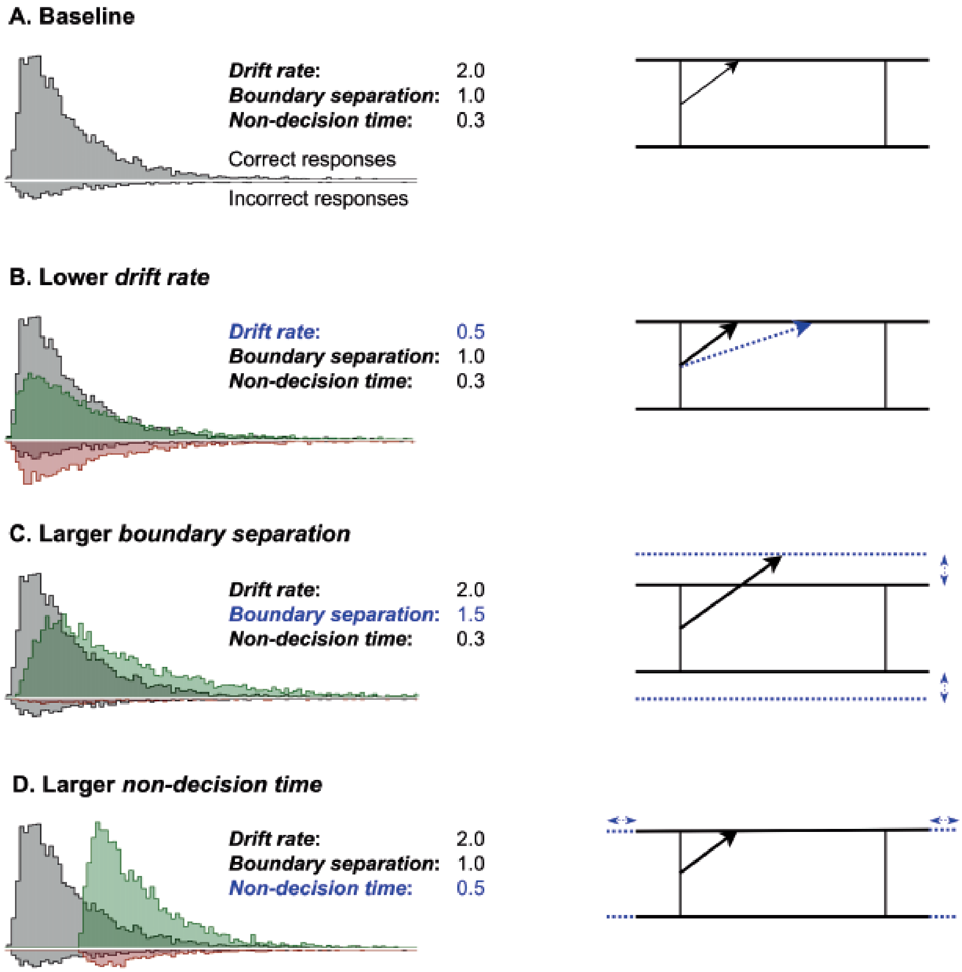



**Figure 1.3.** Simulated effects of changes in Drift Diffusion parameters on response time distributions. Each simulation represents a single person completing 5,000 trials. The upward histograms depict response times of correct responses, and the downward histograms depict response times of incorrect responses. Histograms in grey depict the baseline dataset, and histograms in green/red depict datasets with changes in specific Drift Diffusion parameters. The images on the right graphically depict the change in the corresponding parameter value (in blue). Panel A: Depicts the baseline model, with response times and error rates simulated based on a drift rate of 2.0, a boundary separation of 1.0, and a non-decision time of 0.3. Panel B: A lower drift rate increases the spread in the tail of the distribution but barely changes the peak of the distribution, while increasing error rates. Panel C: A larger boundary separation increases the spread of the distribution and shifts the peak to the right, while decreasing error rates. Panel D: A longer non-decision time shifts the distribution to the right without changing the spread of the distribution and without changing the error rate.

In this dissertation, I focus on the DDM for three reasons. First, previous work shows that the DDM is remarkably flexible. While it was originally developed for simple and fast perceptual and recall tasks, recent work has applied it to a wide range of more complex tasks with longer response windows, such as intelligence and EF tasks (Lerche et al., 2020; Löffler et al., 2024). Second, many established EF tasks have a binary response format and therefore adhere to the key assumptions of the model. Third, the DDM is among the most well-established models in its class, with many recent advances in software and computational approaches that make it increasingly accessible for researchers from fields other than mathematical and cognitive psychology (e.g., D. J. Johnson et al., 2017; Vandekerckhove et al., 2011; Voss et al., 2013, 2015; Wiecki et al., 2013; for some DDM applications in developmental and clinical contexts, see Grange & Rydon-Grange, 2022; Thompson & Steinbeis, 2021).

## 1.4 Specific abilities and general processes

In the context of EF tasks, drift rates can capture specific executive functioning abilities (Löffler et al., 2024). The reason is that a higher EF ability should lead to faster and more accurate responses. For instance, on inhibition tasks like the Flanker task, a person with a higher ability to ignore distractions would be faster at narrowing down attention to goal-relevant information, and would be less likely to accidentally act on distractions. As can be seen in Figure 1.3B, these are the exact response patterns that are associated with an increased drift rate.

However, recent research shows that in addition to specific EF abilities, drift rates on EF tasks also reflect task-general processing speed (Hedge et al., 2022; Lerche et al., 2020; Löffler et al., 2024; Weigard et al., 2021). While EF abilities are specific to particular tasks, task-general processing speed affects performance across EF tasks. The relative contributions of task-general processing speed and specific EF abilities to drift rates can be teased apart using structural equation modeling. Specifically, task-general processing speed can be captured by a task-general latent factor loading on drift rates of all tasks, and specific EF abilities can be captured using task- or ability-specific latent factors of drift rates (Figure 4). After accounting for task-general processing speed, remaining variance should be a more precise measure of specific EF abilities.

Research applying structural equation modeling to drift rates on EF tasks shows that drift rates consistently form a task-general factor that accounts for a substantial part of the variance (Frischkorn et al., 2019; Hedge et al., 2022; Löffler et al., 2024; Weigard & Sripada, 2021). In fact, several studies did not find any meaningful variance associated with specific EF abilities after accounting for task-general processing speed (Frischkorn et al., 2019; Hedge et al., 2022; Löffler et al., 2024), although one study found a correlation between task-general drift rate and self-reported self control, which is related to EF (Weigard et al., 2021). Thus, it remains an open question to what extent traditional EF tasks are suitable measures of specific EF abilities, even when using cognitive modeling.

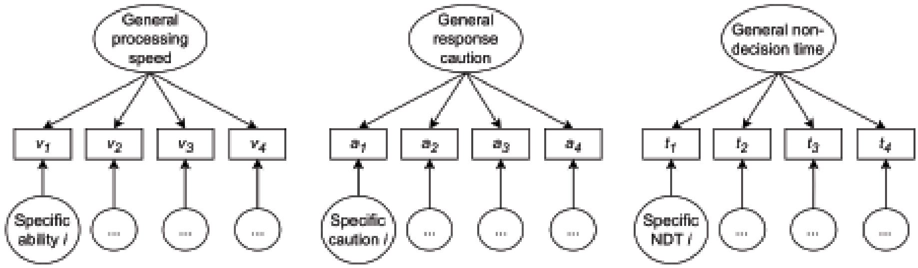



**Figure 1.4.** Task-specific and task-general aspects of performance on EF tasks. Rectangles represent Drift Diffusion parameter estimates on (hypothetical) individual tasks. The ellipses at the top depict general factors that account for shared variance across tasks. The ellipses at the bottom depict task-specific factors that capture residual variances, i.e., the proportion of variance in Drift Diffusion parameters unique to a particular task after partialling out task-general variance. $v$ = Drift rate, $a$ = Boundary separation, $t$ = Non-decision time.

The influence of task-general processing speed on EF task performance is another potential source of bias in adversity research. If adversity is negatively associated with lower task-general processing speed, this could make it seem as if several different EF abilities are impaired, rather than one general process. This has important implications not just for basic science, but also for interventions. Specifically, if adversity is associated with general rather than specific processes, then interventions targeting specific abilities (e.g., training performance on inhibition tasks) may not be effective.

## 1.5 Open questions for adversity research

The integration of deficit and adaptation frameworks is hindered by relying on the use of raw performance scores (see section 1.2). In particular, the idea that performance on EF tasks is influenced by multiple processes has two important implications for adversity research. First, both deficits and adaptations could affect different processes on the same EF task. That is, raw performance on a single task could be lowered by a deficit in one process, while also being enhanced by an adaptation in another process. Second, slower task-general processes, such as basic processing speed, could make it seem as though adversity exposure lowers many different abilities, rather than a single general process. Such task-general effects could also overshadow adaptations in specific abilities, making it difficult to discover unique strengths (Bignardi et al., 2024; Young et al., 2024). Both limitations stand in the way of a full integration of deficit and adaptation frameworks, and common practices based on raw performance scores fall short on both counts.

Revisiting Yara's case, a cognitive modeling analysis of her performance might reveal that her slower responses result from anxiety about performing the tasks, prompting her to be cautious in order to avoid making mistakes. In addition, her tendency to be slower across tasks may be caused by slower general processing speed. After accounting for these factors, we may discover that Yara's difficulties with inhibition and working memory specifically are smaller than initially thought. We may even find that she has enhanced or intact abilities that remained hidden before. This would have important consequences for the nature of interventions. Instead of cognitive training exercises and behavioral strategies focused on improving specific abilities, interventions could instead focus on finding and alleviating the source of her task anxiety, and allowing her to work on tasks at her own pace. Thus, solving these methodological challenges has important implications not just for basic science, but also for interventions.

## 1.6 Current aims

This dissertation has three central aims. The first aim is to uncover the cognitive processes underlying performance differences (both lowered and enhanced) in people exposed to adversity. Using a combination of DDM and structural equation modeling, I will show that researchers likely overestimate deficits in specific cognitive abilities when analyzing raw performance alone. The second aim is to investigate to what extent performance differences on EF tasks can be attributed to ability-specific processes as opposed to more general processes or strategies. The third aim is to show how moving beyond raw performance towards cognitive processes can enrich the next generation of adversity research.

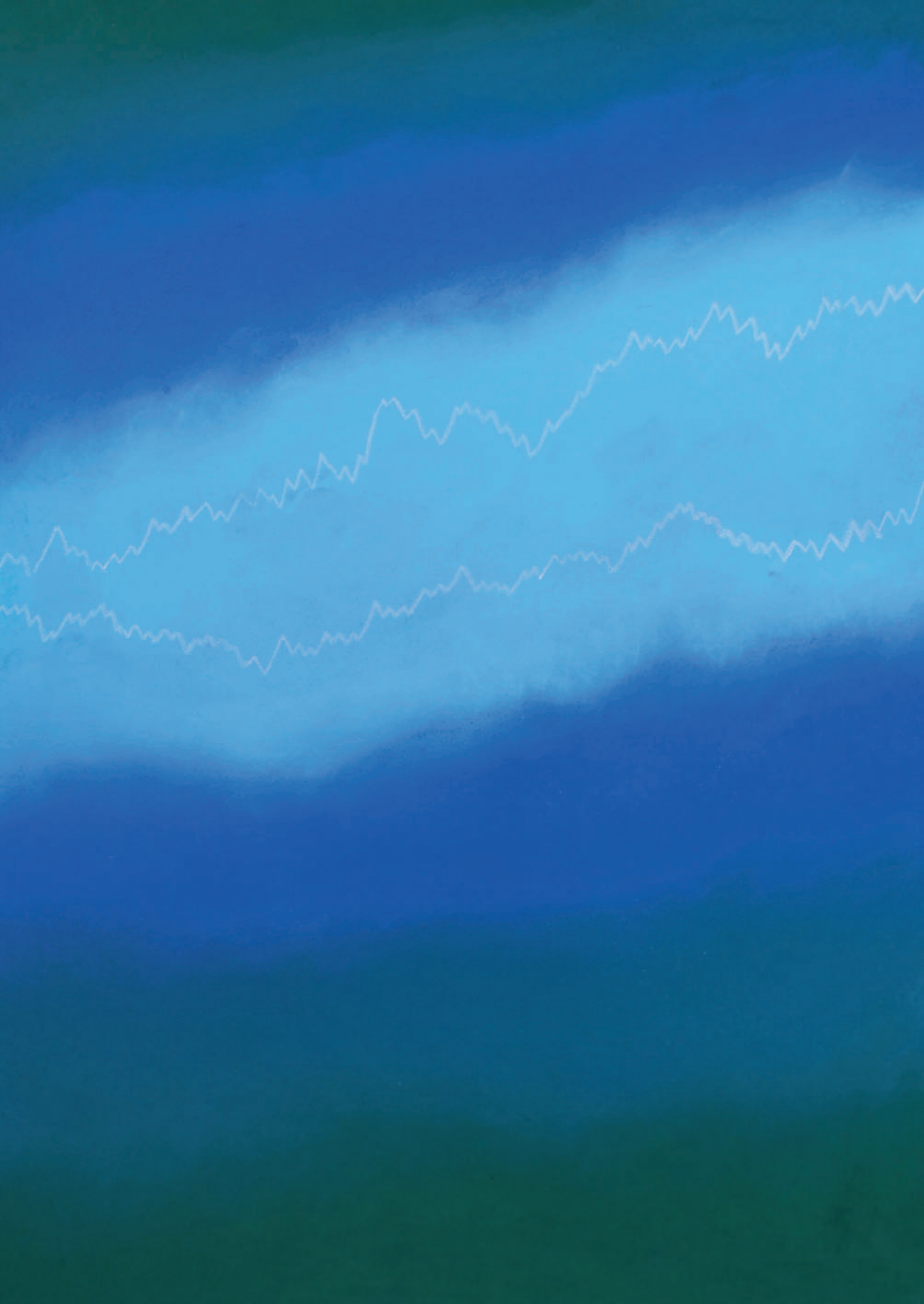
## 1.7 Dissertation outline

The chapters in this dissertation can be read in any order. The empirical chapters (Chapters 2-5) are based on articles that have either been published in or submitted to peer-reviewed scientific journals. In **Chapter 2**, I analyze the associations of two forms of adversity—material deprivation and household threat—with inhibition ability, attention shifting ability, mental rotation ability, and general processing speed, among a representative sample of children from the United States. Specifically, I use DDM and structural equation modeling to investigate which cognitive processes are associated with adversity in 9-10 year-olds, and whether these associations are more task-general or task-specific. In **Chapter 3**, I analyze the associations of two forms of adversity —exposure to material deprivation and threat—with inhibition ability, attention shifting, and general processing speed, among a representative sample of adults from the Netherlands. As in Chapter 2, I use DDM and structural equation modeling, but this time including two inhibition tasks, three attention shifting tasks, and a basic processing speed task, in order to estimate more precise latent ability factors. In **Chapter 4**, I analyze the associations of two forms of adversity—exposure to violence and unpredictability—with inhibition ability, among young adults from the United States. Across three studies, I use the Shrinking Spotlight Model—a special version of the DDM—which captures attention processes related to inhibition. In **Chapter 5**, I analyze the associations of three forms of adversity—exposure to neighborhood threat, material deprivation, and unpredictability—with working memory ability, among a representative sample of adolescents and adults from the Netherlands. Specifically, I use structural equation modeling to distinguish between working memory capacity and working memory updating. In **Chapter 6**, I discuss the insights generated in Chapters 2-5, and develop a roadmap for future adversity research in the context of two developments in the field: integrating deficit and adaptation frameworks, and developing more ecologically and contextually relevant measurement instruments of EF.

## 1.8 Open science statement

For all empirical chapters (Chapters 2-5), I preregistered the hypotheses, design, and analyses prior to collecting the data and/or conducting the analyses. All deviations from preregistrations are described in the main text. The studies reported in Chapter 2 and Chapter 5 were Registered Reports. In a Registered Report, the Introduction and Methods sections are submitted to and peer-reviewed by the journal prior to data collection and/or analyzing the data (Chambers & Tzavella, 2021). The Registered Report described in Chapter 5 was peer-reviewed through *Peer Community In Registered Reports*, a non-commercial initiative that offers peer review of preprints outside of traditional journals (see https://rr.peercommunityin.org/PCIRegisteredReports).

For each empirical chapter, the analysis code, study materials, (synthetic) data, and reproducible manuscript are openly available on my personal GitHub page (https://github.com/stefanvermeent). Each chapter provides links to the respective GitHub repositories, which were turned into user-friendly website versions. Full project histories with timestamped milestones were generated using the *projectlog* R package (Vermeent, 2023), which I developed in an attempt to optimize my Open Science workflow. Chapter 2 is based on data from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), and for that reason cannot be shared openly on the Github Repository. The same is true for Chapter 3 and 5, which are based on a combination of previously collected and newly collected data from the Longitudinal Internet Studies for the Social Sciences (LISS) panel study (https://lissdata.org). Researchers with an academic affiliation can apply for access to both data sets.



# Chapter 2

## Cognitive deficits and enhancements in youth from adverse conditions: An integrative assessment using Drift Diffusion Modeling in the ABCD study

**This chapter is based on**
Vermeent, S., Young, E.S., DeJoseph, M.L., Schubert, A.-L., & Frankenhuis, W.E. (2024). Cognitive deficits and enhancements in youth from adverse conditions: An integrative assessment using Drift Diffusion Modeling in the ABCD study. *Developmental Science, 27*(4), e13478. https://doi.org/10.1111/desc.13478

## 2.0 Abstract

Childhood adversity can lead to cognitive deficits or enhancements, depending on many factors. Though progress has been made, two challenges prevent us from integrating and better understanding these patterns. First, studies commonly use and interpret raw performance differences, such as response times, which conflate different stages of cognitive processing. Second, most studies either isolate or aggregate abilities, obscuring the degree to which individual differences reflect task-general (shared) or task-specific (unique) processes. We addressed these challenges using Drift Diffusion Modeling (DDM) and structural equation modeling (SEM). Leveraging a large, representative sample of 9-10 year-olds from the Adolescent Brain Cognitive Development (ABCD) study, we examined how two forms of adversity—material deprivation and household threat—were associated with performance on tasks measuring processing speed, inhibition, attention shifting, and mental rotation. Using DDM, we decomposed performance on each task into three distinct stages of processing: speed of information uptake, response caution, and stimulus encoding/response execution. Using SEM, we isolated task-general and task-specific variances in each processing stage and estimated their associations with the two forms of adversity. Youth with more exposure to household threat (but not material deprivation) showed slower task-general processing speed, but showed intact task-specific abilities. In addition, youth with more exposure to household threat tended to respond more cautiously in general. These findings suggest that traditional assessments might overestimate the extent to which childhood adversity reduces specific abilities. By combining DDM and SEM approaches, we can develop a more nuanced understanding of how adversity affects different aspects of youth's cognitive performance.

**Author contributions**

All authors were involved in conceptualizing the study. SV accessed and analyzed the data, and wrote the first draft of the manuscript. All authors provided feedback on the manuscript.

## 2.1 Introduction

The effects of early-life adversity—such as growing up in poverty or experiencing high levels of violence—on cognition are complex. There is a growing consensus that adversity-exposed youth may develop not only deficits, but also strengths. For example, studies find lowered and improved performance across different cognitive domains including (but not limited to) executive functioning, social cognition, language, and emotion (Ellis et al., 2022; Frankenhuis et al., 2016; Frankenhuis & Weerth, 2013; Sheridan et al., 2022; Sheridan & McLaughlin, 2014). Researchers focused on one type of effect or another acknowledge the importance of identifying both deficits and strengths. Yet, in practice, they often focus on one at the expense of the other. To develop an integrated, well-rounded, and nuanced understanding of how adversity shapes cognitive abilities, research must integrate both types of effects.

Such an integration of deficit- and strength-based approaches is hampered by two methodological challenges. First, most cognitive tasks involve different stages of processing which are obscured when analyzing raw performance differences. This makes it difficult to understand why cognitive performance may be lowered or improved. Second, adversity may lower or improve performance because it affects general processes (i.e., processes shared across many tasks) or abilities that are task-specific. In this Registered Report, we use a framework that tackles both challenges. First, we decompose raw performance into measures of different stages of cognitive processes through cognitive modeling. Second, we analyze four different tasks—tapping processing speed, attention shifting, inhibition, and mental rotation—all of which have documented associations with adversity. Finally, we model shared (i.e., task-general) and unique (i.e., task-specific) factors that drive performance and investigate how they are associated with adversity.

### What do deficit and enhancement patterns mean?

Both the deficit and strength-based literature often use speeded tasks, in which participants are usually instructed to respond as fast and accurate as possible. For example, performing well on inhibition tasks (e.g., Flanker task, Go/No-Go Task; Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Noble et al., 2005), attention shifting tasks (e.g., Dimensional Change Card Sort; Farah et al., 2006; Fields et al., 2021; Mittal et al., 2015; Noble et al., 2005; Nweze et al., 2021; Young et al., 2022), and stimulus detection tasks (Farah et al., 2006; Noble et al., 2005; Pollak, 2008) requires fast and accurate responses. In practice, performance is often quantified using aggregated indices of speed alone (e.g., RT), accuracy alone (e.g., proportion correct), or both independently (rather than in an integrated manner).

In both the deficit and strength-based literature, *task performance* (indexed by mean RTs or accuracy) is routinely equated with *cognitive ability*. For example, deficit-focused studies relate slower RTs on inhibition tasks to *worse inhibition ability* (Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Noble et al., 2005). Strength-based studies relate faster RTs on standard attention shifting tasks to *better shifting ability* (Fields et al., 2021; Mittal et al., 2015; Young et al., 2022). However, speed and accuracy comprise more than pure ability (e.g., inhibition, attention shifting). They also measure other constructs such as response caution (e.g., more or less cautious responding), speed of task preparation (e.g., orienting attention, encoding information), and speed of response execution. This heterogeneity creates an inferential risk, namely, if performance differences are interpreted as differences in abilities without sufficiently considering alternative explanations. In addition, the effect of adversity exposure may not be limited to a single process. For example, a specific type of adversity could affect both the speed of information processing and also shape the strategy that a person uses. These inferential challenges have real-world implications, especially when raw performance is used as an early screening tool to assess cognitive abilities (Distefano et al., 2021).

One promising solution to these issues is leveraging cognitive measurement models developed by mathematical psychologists. For speeded binary decision tasks, a well-established measurement model is the Drift Diffusion Model (DDM; Forstmann et al., 2016; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Wagenmakers, 2009). The DDM integrates speed and accuracy on a trial-by-trial level to estimate cognitive processes at different stages of the decision-making process. The DDM assumes that people go through three distinct phases on each trial (see Figure 2.1 for a visualization). The first phase, *preparation*, includes processes such as focusing attention and visually encoding the stimulus. In the second phase, *decision-making*, people gather evidence for both response options until the evidence sufficiently favors one option over the other (explained below) and the decision process terminates. The third phase, *execution*, involves preparing and executing the motor response corresponding to the choice.

DDM estimates a set of parameters[3] for each participant that represent each phase of the decision process (Voss et al., 2004). The *drift rate* (*v*) represents the speed of information uptake (Schmiedek et al., 2007; Voss et al., 2013). People with a higher drift rate are faster and make fewer errors. The *non-decision time* (*t0*) includes initial preparatory processes (e.g., visually encoding the stimulus) and processes after the decision is made (e.g., pressing a button). All else being equal, longer non-decision times reflect slower information processing but without a cost nor benefit in accuracy. The *boundary separation* (*a*) represents the distance between the two decision boundaries. A larger boundary separation means more information is collected before making a decision. Thus, boundary separation measures response caution. In contrast to non-decision time, larger boundary separation leads to slower but more accurate responses, reflecting a speed-accuracy tradeoff.

As mentioned earlier, adversity-related raw performance differences—both lowered and improved performance—are typically interpreted as differences in ability

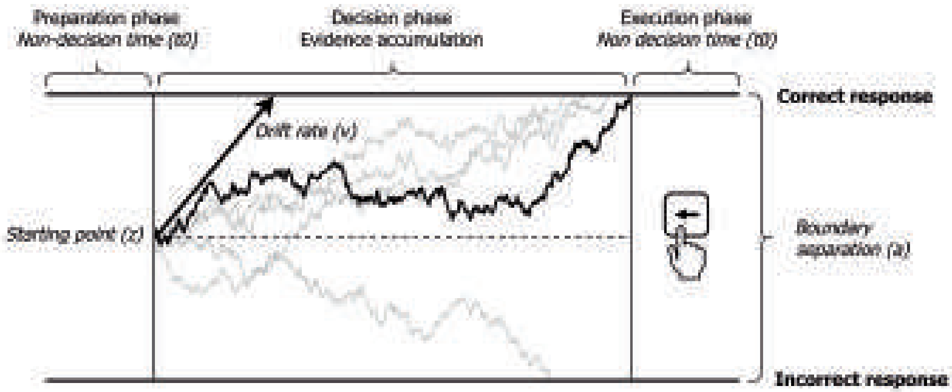



**Figure 2.1.** A visual overview of the Drift Diffusion Model (DDM). The DDM assumes that decision making on cognitive tasks with two forced response options advances through three stages. First, people go through a preparation phase in which they engage in initial stimulus encoding. Second, people gather information for one of two response options until the accumulation process terminates at one of the decision boundaries. Each squiggly line represents the evidence accumulation process on a single trial. Third, a motor response is triggered in the execution phase. The model estimates four parameters that reflect distinct cognitive processes (printed in italic): (1) The *drift rate* represents the rate at which evidence accumulation drifts towards the decision boundary and is a measure of processing speed; (2) The *non-decision time* represents the combined time spent on task preparation and response execution; (3) The *boundary separation* represents the width of the decision boundaries and is a measure of response caution; (4) The *starting point* represents the starting point of the decision process and can be used to model response biases (not considered in this study).

(e.g., inhibition, attention shifting). If these interpretations are accurate, then drift rate would reflect these variations. That is because improved ability would result in both decreased RTs and increased accuracy. However, if performance differences arise through other factors—such as differences in response caution or response speed—they would be captured by parameters other than the drift rate. Thus, disentangling the drift rate, non-decision time, and boundary separation enhances our understanding of how adversity exposure is associated with performance.

### Are deficit and enhancement patterns task-specific or task-general?

An important caveat to interpreting task performance on any task in isolation is that performance on most tasks relies both on shared cognitive processes and unique abilities. For example, RTs on executive functioning tasks are substantially confounded with general processing efficiency (Frischkorn et al., 2019; Lerche et al., 2020; Löffler et al., 2024). Both task-specific abilities and task-general processes affect RTs and accuracy in similar ways and are thus likely confounded in drift rates. Task-general effects create the illusion that many different abilities are affected by adversity when in fact only one more general process is affected. Consider research on cognitive deficits. Adversity exposure might disrupt general cognitive processes shared across many tasks, such as general processing speed, for example, because of its effects on brain regions that are involved across several cognitive abilities (Sheridan & McLaughlin, 2014). If so, studies analyzing raw Flanker performance in isolation will find processing speed deficits but wrongly interpret this as an inhibition deficit. Such distinctions matter for both deficit- and strength-based approaches (e.g., does adversity impair broad domains such as executive functioning? Does it enhance specific abilities such as attention shifting?), as well as for real-world interventions grounded in both approaches (e.g., school-based interventions targeting broad domains or specific abilities).

Structural equation modeling (SEM) can disentangle task-general and task-specific processes. For example, it can estimate shared task variance with latent task-general variables. By estimating shared variance across different tasks, we can also obtain more precise estimates of task-specific abilities (i.e., variance unique to specific tasks). Bignardi et al. (2024) recently applied this approach to model how socioeconomic status (SES) is related to standard performance measures in three large data sets. They used SEM to model the effect of SES on a general factor and task-specific residual variances. Lower SES was associated with a lower general ability, but *enhanced* task-specific processing speed, inhibition, and attention shifting. However, their analysis looked at shared and unique variance using raw performance measures. Thus, it is subject to the same limitations outlined in the previous section.

**The current study**

Here, we analyzed the Adolescent Brain Cognitive Development (ABCD) study data (http://abcdstudy.org). The ABCD study is ideal because it provides a large, representative, and socioeconomically and ethnically diverse sample of 9- to 10 year-olds—an age range characterized by rapid growth in cognitive abilities (Blakemore & Choudhury, 2006).

We studied two dimensions of adversity: household threat and material deprivation. These forms of adversity have been widely studied in their relation to cognitive outcomes—from both deficit and strength-based perspectives (Fields et al., 2021; Schäfer et al., 2022; Sheridan et al., 2022; Young et al., 2022)—and are central to contemporary conceptualizations of adversity (e.g., McLaughlin et al., 2021; Sheridan & McLaughlin, 2014). Prior work has shown that cognitive deprivation is more strongly associated with lower cognitive performance than threat exposure (Salhi et al., 2021; Sheridan et al., 2020). Although material deprivation (as measured here) and cognitive deprivation (in previous studies) are not identical, both seem related to access to resources that support cognitive development (e.g., books in the home, formal education). Indeed, in the ABCD sample material deprivation is highly or moderately correlated with income (-.81) and education (-.56), while correlations with household threat are lower (-.25 and -.12, respectively; DeJoseph et al., 2022). Therefore, to the extent that the deprivation-versus-threat literature has captured ability-relevant processes, we may expect material deprivation to be more strongly associated with lower drift rates than threat exposure.

We analyzed four cognitive abilities that have been studied in relation to adversity. We included *attention shifting* because previous work has reported enhancement of this ability in children and (young) adults with more exposure to environmental unpredictability (based on raw performance switch costs; Fields et al., 2021; Mittal et al., 2015; Young et al., 2022; but see Nweze et al., 2021). Theoretically, attention shifting is thought to enable people to rapidly adjust to, and take advantage of, a changing environment (e.g., seize fleeting opportunities). We included *inhibition* because previous research suggests that children with more adverse experiences are worse at inhibiting distracting information (based on raw RT difference scores; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015; Tibu et al., 2016). We included *mental rotation* because previous studies have found negative associations between SES and mental rotation ability (based on RTs and accuracy; Assari, 2020; Bignardi et al., 2024). To the extent that these performance differences reflect differences in the respective abilities—as they have been interpreted—they should show up in *task-specific drift rates*. We also included a measure of *processing speed*, which was not measured in relation to adversity but provided a direct measure of the type of basic processing speed that plays a role in the other tasks. Taken together, the four tasks provided a broad assessment of cognitive domains, which makes them well-suited for isolating task-general processes. As all four tasks adhere to DDM assumptions, we could compare them based on the same model parameters.

Adaptation-based frameworks predict increased task-specific drift rates. This follows from the key assumption that adversity shapes specific abilities, rather than general cognitive processes (Ellis et al., 2022; Frankenhuis et al., 2016; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013). Task-specific enhancement in the attention-shifting drift rate would align with this assumption, as this ability is thought to be adaptive in changing environments; but enhancement in the task-general drift rate would not. One study reports evidence suggesting that exposure to threat but not deprivation is associated with better attention shifting (Young et al., 2022). If so, we should expect to see higher task-specific drift rates with household threat, but not with material deprivation. Enhanced task-specific drift rates on inhibition and mental rotation would be unexpected yet interesting. It would constitute novel documentation of enhancements, and would suggest that lowered raw performance reflects ability-irrelevant processes. Finally, equivalent drift rates across adversity levels would also not be consistent with strength-based frameworks; rather, such a pattern would suggest that abilities are intact (i.e., not affected by adversity).

Deficit perspectives can accommodate both lowered task-specific and lowered task-general drift rates. On the one hand, past work suggests that adversity impairs specific abilities (e.g., inhibition; Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015). On the other hand, there is also evidence that adversity affects general cognitive ability (Bignardi et al., 2024)—perhaps through its broad effects on brain regions that are involved across several cognitive abilities (Sheridan & McLaughlin, 2014). However, equivalent or enhanced drift rates, whether they be task-specific or task-general, would not be consistent with deficit perspectives; rather, this would suggest that abilities are intact or enhanced.

Our approach adds value in a third way besides separating drift rate from ability-irrelevant factors and isolating task-specific and task-general effects: It allows us to quantify cognitive deficits and enhancements separately within the same model. This is because the task-specific and task-general estimates are statistically independent. Thus, for instance, we may find that adversity lowers general drift rate, as well as some task-specific drift rate (e.g., capturing inhibition), but increases other task-specific drift rates (e.g., capturing attention shifting).

If the drift rates we observe align with previous interpretations of performance differences as outlined above, our findings support existing theories about deficits and enhancements. However, if not drift rates, but non-decision time or boundary separation account for the existing findings, and drift rates do not, neither deficit- or adaptation-based frameworks are supported. This would at a minimum invite reflection—perhaps revision—of the evidence base for (parts of) these frameworks. At the same time, such findings would offer clear directions for future research in this field (e.g., which factors explain variation in non-decision times and/or boundary separation across levels of adversity). Thus, regardless of the specific pattern of outcomes, our analyses contribute to an accurate and refined understanding of how early-life adversity shapes cognitive abilities.

## 2.2 Methods

### Sample

The ABCD study (http://abcdstudy.org), is a prospective, longitudinal study of approximately 12,000 youth across the United States. We focused on the baseline assessment, which has the largest collection of cognitive tasks suitable for DDM (Luciana et al., 2018). There were four tasks: (1) Processing Speed Task (Pattern Comparison Processing Speed Task), (2) Attention Shifting Task (Dimensional Change Card Sort Task), (3) Inhibition Task (Flanker Task), and (4) Mental Rotation Task (Little Man Task). At baseline, the study included 11,878 youths (aged between 9 and 10 years old, measured in months) recruited across 21 sites. The study used multi-stage probability sampling to obtain a nationally representative sample (Heeringa et al., 2010). Baseline assessments were completed between September 1st 2016 and August 31st 2018 (see Garavan et al., 2018). Our analysis sample includes 10,687 participants who had trial-level data available on all four[4] cognitive tasks.

## Open Science Statement

All analysis scripts, materials, and instructions needed to reproduce the findings are available on the article's Github repository (https://stefanvermeent.github.io/abcd_ddm/). The raw study data cannot be shared on public repositories. Personal access to the ABCD dataset is required to fully reproduce our analyses and can be requested at https://nda.nih.gov.

We obtained access to the full ABCD data repository and performed initial data cleaning and analyses *prior* to Stage 1 submission. However, we preprocessed cognitive task data in isolation to prevent biasing the analyses involving independent variables. The goal of these analyses was to assure that the pre-selected cognitive tasks adhered to basic DDM assumptions and had the required trial-level data available in the right format. These initial analyses were preregistered (https://stefanvermeent.github.io/abcd_ddm/preregistrations/README.html).

To increase transparency, we developed an automated workflow (using R and Git) to track the data files read into the analysis environment. First-time access to any data file was automatically tracked via Git, providing an overview including the timestamp, a description of the data, and the R code that was used to read in the data. The supplemental materials provide a detailed description and visual overview of this workflow. An overview of the data access history is provided in the repository's README file (https://stefanvermeent.github.io/abcd_ddm/).

## Exclusion Criteria

For the cognitive task data, we applied exclusion criteria in two steps: first, cleaning trial-level data, and second, removing participants with problematic trial-level data (discussed below). For both, most criteria were as preregistered, but a few deviated from or were additional to the preregistration. The data processing steps described below were preregistered unless noted otherwise.

First, we removed RTs of the Attention Shifting, Flanker, and Mental Rotation Tasks that exceeded maximum task-specific RT thresholds (> 10 seconds (0.07%), > 10 seconds (0.04%), and > 5 seconds (< 0.01% of trials), respectively). The Processing Speed Task did not have a programmed time-out. Instead, we cut off responses > 10 seconds (0.15% of trials) to remove extreme outliers. This step was not preregistered as we did not anticipate these extreme outliers.

Next, we removed trials with: (1) RTs < 300 ms (ranging from 0.01% to 1.03% of trials across tasks); (2) RTs > 3 *SD* above the participant-level average log-transformed mean RT (ranging from 0.02% to 0.85% of trials across tasks; the same thing was done for RTs < 3 SD on the Processing Speed Task (not preregistered) to remove several fast outliers); (3) trials with missing response times and/or accuracy data (< 0.01% for all tasks except Mental Rotation). We found that the response time-out of 5 seconds on the Mental Rotation Task led to missing responses on 10.55% of trials. This truncated the right-hand tail of the RT distribution, which can bias DDM estimation. Therefore, we decided to impute these values during DDM estimation instead of removing them (see the Supplemental materials for more information).

Next, we excluded participants who (1) had suffered possible mild traumatic brain injury or worse (*n* = 118); (2) showed a response bias of > 80% on a task (ranging between zero and 212; deviating from the preregistration); (3) had a low number of trials left after trial-level exclusions, defined as < 20 trials for Mental Rotation and Attention Shifting (*n* = zero and 19, respectively) and < 15 trials for Flanker and Processing Speed (*n* = 64 and 34, respectively, deviating from the preregistration). Finally, we excluded task data of several participants based on data inspection (not preregistered): two participant with 0% accuracy on the Mental Rotation Task; two participants who showed a sharp decline in accuracy over time on the Processing Speed Task; 49 participants on the Attention Shifting Task who (almost) only made switches across all trials, even on repeat trials. We also decided to include participants with missing data on one or more tasks because our main analyses used FIML for missing data.

The final sample consisted of 10,563 participants (See Table 2.1).

**Table 2.1.** Descriptive statistics for the training and test set.

| | Training | Test |
|---|---|---|
| N | 1500 | 9063 |
| Sex (%) | 48.7 | 47.6 |
| Age (Mean (SD)) | 119 (7.5) | 119 (7.4) |
| Parent highest education in years (Mean (SD)) | 20.3 (2.4) | 20.3 (2.4) |
| Race | | |
| White (%) | 53.5 | 55.6 |
| Black (%) | 16.6 | 16.1 |
| Hispanic (%) | 16.9 | 15.6 |
| Other (mixed, Asian, AIAN, NHPI) | 12.9 | 12.7 |
| Income-to-needs (Mean (SD)) | 3.8 (2.4) | 3.7 (2.4) |

## Measures

### Cognitive Tasks

**Inhibition Task.** The NIH Toolbox Flanker task is a measure of cognitive control and attention (Zelazo et al., 2014). On each trial, participants saw five arrows that were positioned side-by-side. The four flanking arrows always pointed in the same direction, either left or right. The central arrow either pointed in the same direction (congruent trials) or in the opposite direction (incongruent trials). Participants were instructed to always ignore the flanking arrows and to indicate whether the central arrow is pointing left or right. After four practice trials, participants completed 20 test trials, of which 12 were congruent ($Mean_{RT}$ = 0.84 seconds, $SD$ = 0.28) and eight were incongruent ($Mean_{RT}$

= 1.02 seconds, *SD* = 0.44). The standard outcome measure is a normative composite of accuracy and RT. For more information on the exact calculation, see Slotkin et al. (2012).

**Processing Speed Task.** The NIH Toolbox Pattern Comparison Processing Speed task (Carlozzi et al., 2015) is a measure of visual processing. On each trial, participants saw two images and judged whether the images were the same or different. When images were different, they varied on one of three dimensions: color, adding or taking something away, or containing more or less of a particular item. The standard outcome measure is the number of items answered correctly in 90 seconds (normalized). On average, participants completed 38.96 trials ($Mean_{RT}$ = 2.24 seconds, *SD* = 0.47).

**Attention Shifting Task.** The NIH Toolbox Dimensional Change Card Sort Task is a measure of attention shifting or cognitive flexibility (Zelazo, 2006; Zelazo et al., 2014). A white rabbit and green boat were presented at the bottom of the screen. Participants matched a third object to the rabbit or boat based on either color or shape. After eight practice trials, participants completed 30 test trials alternating between shape and color in pseudo-random order. Of these, 23 were *repeat* trials (i.e., the sorting rule was the same as on the previous trial; $Mean_{RT}$ = 1 seconds, *SD* = 0.36) and 7 were *switch* trials (i.e., the sorting rule was different than on the previous trial; $Mean_{RT}$ = 1.03 seconds, *SD* = 0.39). The standard outcome measure is a normative composite of accuracy and RT. For more information on the exact calculation, see Slotkin et al. (2012).

**Mental Rotation Task.** The Little Man task (referred to in this article as the Mental Rotation task) is a measure of visual-spatial processing (Luciana et al., 2018). Participants saw a simple picture of a male figure holding a briefcase in his left or right hand. They had to indicate whether the briefcase was in the left or right hand. The image could have one of four orientations: right side up or upside down, and facing towards or away from the participant. Thus, on half of the trials, participants had to mentally rotate the image in order to make the decision. Participants first completed three practice trials and then completed 32 test trials ($Mean_{RT}$ = 2.65, *SD* = 0.47). The standard outcome measure is an efficiency measure, calculated as the percentage correct divided by the average RT.

### Adversity measures

**Material deprivation.** We assessed material deprivation with seven items from the parent-reported ABCD Demographics Questionnaire. These items originate from the Parent-Reported Financial Adversity Questionnaire (Diemer et al., 2013). The items assess whether or not (1 = Yes, 0 = No) the youth's family experienced several economic hardships over the 12 months prior to the assessment (e.g., 'Needed food but couldn't afford to buy it or couldn't afford to go out to get it').

We used a previously created factor score of this measure derived from MNLFA (Bauer, 2017). This score empirically adjusts for measurement non-invariance across sociodemographic characteristics and creates person-specific factor scores that enhance measurement precision and individual variation (Curran et al., 2014). In short, MNLFA scores assume a common scale of measurement across groups and age, as well as adjust for measurement biases that would have otherwise biased our substantive analyses. DeJoseph et al. (2022) describe how this score was computed. Higher scores indicate more material deprivation.

**Household threat.** We assessed threat experienced in the youth's home using the Family Conflict subscale of the ABCD Family Environment Scale (Moos, 1994; Zucker et al., 2018). The subscale consisted of nine items assessing conflict with family members (e.g., 'We fight a lot in our family'). Items were endorsed with either 1 (True) or 0 (False). Two items were positively valenced and were therefore reverse-scored. Similar to material deprivation, we used a previously-created factor score of this measure derived from MNLFA (DeJoseph et al., 2022). Higher scores indicate more threat exposure.

**Sociodemographic covariates.** Several sociodemographic covariates were included in the SEM models (see Planned Main Analyses) that use the MNLFA scores representing material deprivation and household threat exposure. This is because MNLFA scores are adjusted for these covariates. Thus, it is recommended that variation in these covariates is also adjusted for in dependent variables (Bauer, 2017).

We calculated income-to-needs ratios by first taking the average of each binned income (< \$5000, \$5,000–\$11,999, \$12,000–\$15,999, \$16,000–\$24,999, \$25,000–\$34,999, \$35,000–\$49,999, \$50,000–\$74,999, \$75,000–\$99,999, \$100,000–\$199,999, ≥ \$200,000) as a rough approximation of the family's total reported income. Then we divided income by the federal poverty threshold for the year at which a family was interviewed (range = \$12,486–\$50,681), adjusted for the number of persons in the home. We used highest education (in years) out of the two caregivers (or one if a second caregiver was not provided) as a continuous variable. We collapsed youth race into 4 levels (White, Black, Hispanic, Other) and subsequently dummy-coded with White (the most numerous racial group) serving as the reference category in all models. We dichotomized youth sex such that 1 = Female and 0 = Male. We used youth age (in months) as a continuous variable and centered on the mean.

## Analysis Pipeline

### Primary analyses

The approved Stage 1 Protocol for this manuscript can be found on the Open Science Framework (https://osf.io/4n8qr). Before conducting analyses, we split the full sample up in a training set ($n$ = 1,500) and a test set ($n \approx 8,500$). We conducted our main analyses in three steps (see Figure 2.2): (1) fitting the DDM to the cognitive task data;

(2) fitting the SEM model to the adversity and DDM data and optimize it where necessary based on the training set; (3) Refitting the model to the test data and interpret the regression coefficients. We conducted a simulation-based power analysis based on the main SEM model (see Figure 2.3), with standardized regression coefficients of 0.06, 0.08 and 0.1 and the alpha level set to .05. The analysis indicated that we would have more than 90% power for all regression paths with $N$ between 2,500 ($\beta$ = 0.1) and 6,500 ($\beta$ = 0.06) (see Figure A1.2) .

All analyses were conducted in R 4.2.1 (Team, 2022). The source code can be found on the Github repository (https://stefanvermeent.github.io/abcd_ddm/scripts/README.html).

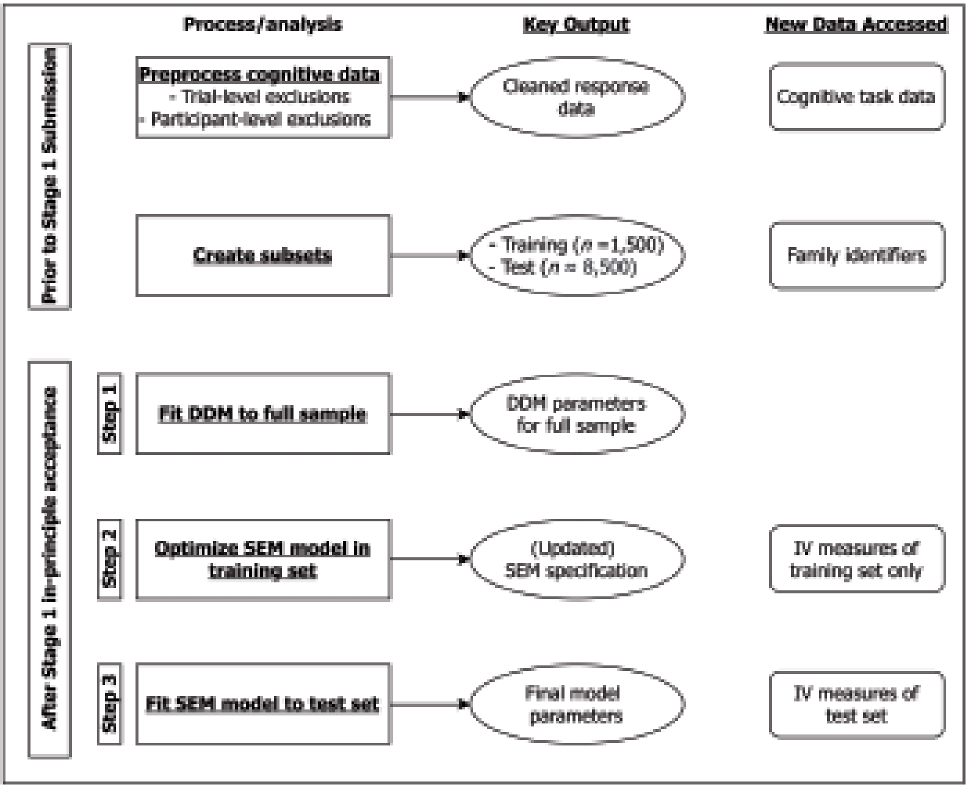



**Figure 2.2.** Visual overview of the full analysis workflow. Analyses were done in two stages: (1) prior to Stage 1 submission of the manuscript, and (2) after Stage 1 in-principle acceptance. Analyses at Stage 1 only focused on the cognitive task data. Independent variables (i.e., threat and deprivation measures) were only accessed during Stage 2 after all DDM models had been fit, and only for the test set after the model had been optimized based on the training set. Data access was tracked via the GitHub repository. IV = independent variable; SEM = structural equation modeling; DDM = Drift Diffusion Model.

**Step 1: DDM estimation.** The DDM was fit to each cognitive task in a hierarchical Bayesian framework which estimates DDM parameters both on the individual and group level (Vandekerckhove et al., 2011; Wiecki et al., 2013). We used code provided by D. J. Johnson et al. (2017). The benefit of this approach is that group-level information is leveraged to estimate individual-level estimates. This differs from classic DDM estimation approaches where the model is fitted to the data of each participant separately (Voss et al., 2013). This is particularly useful in developmental samples like the ABCD dataset which have a limited number of trials per participant but substantially larger sample sizes than is typical in the DDM literature. We ran parameter recovery studies simulating data based on the Inhibition task, which has the lowest overall number of trials. Parameter recovery was excellent for the scenario that we planned in our main analyses (all *r*s ≥ .84). See the supplemental materials for more details.

All models freely estimated the drift rate, non-decision time, and boundary separation while constraining response bias to 0.5 (i.e., assuming no bias towards a particular response option). For the Flanker and Attention Shifting Task, we compared model versions that separately estimate drift rate and non-decision-time per task condition or collapsed across conditions. Boundary separation was constrained to be the same across conditions. For the Processing Speed Task and the Mental Rotation Task, we estimated DDM parameters across all trials. The best-fitting model of each task was used to estimate participant-level DDM parameters. See the supplement for more information about model fitting procedures.

**Step 2: Model optimization in training set.** We first estimated and (where necessary) optimized the SEM in the training set using the *lavaan* package (Rosseel, 2012). The goal of this step was to investigate whether we needed to adjust the model specification in any way (e.g., add residual correlations, introduce or reduce constraints of factor loadings, etc.) to achieve good model fit. For this reason, the model fitted in this step was not interpreted to address our research aims.

See Figure 2.3 for the *a-priori* specification of the model. In the measurement model, all three DDM parameters across all tasks (i.e., drift rates, non-decision times, and boundary separations) loaded on separate latent factors for each parameter type. Unique (residual) variances of the manifest (i.e., measured) DDM parameters were captured in additional latent factors (one per parameter). The structural model estimated regression paths going from each adversity measure (see Adversity measures) to the general latent factors and to the unique variances of the DDM parameters of each task. For model identification reasons, we did not estimate regression paths to the unique variances of the Processing Speed Task. We first estimated and optimized the measurement models separately for each diffusion model parameter, which allowed us to efficiently detect sources of potential badness of fit. Once measurement models provided an adequate account of the data, we integrated them into the structural model shown in Figure 2.3. In addition, clustering of siblings and twins within families was accounted

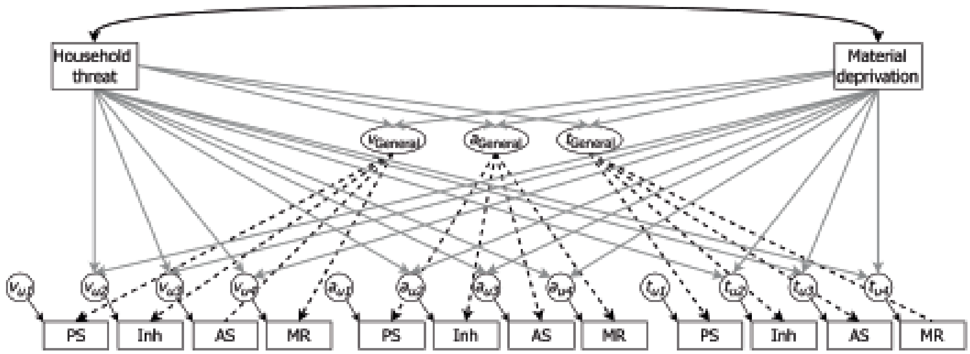



**Figure 2.3.** Visualization of the full structural equation model (SEM). Elipses represent task-general factors. Circles represent task-specific (residual) variances. Dotted black lines represent covariances. Dashed black lines represent factor loadings. Solid grey lines represent regression paths. The factor loadings to each of the Processing Speed Task indicators are fixed to 1. The factor loadings of the task-specific factors are fixed to 1 and the residual variances of the manifest indicators are fixed to 0. For model identification reasons, we do not estimate regression paths to the unique variances of the Processing Speed Task. Not shown in this Figure to improve readability: (1) the sociodemographic covariates that are included in the MNLFA scores (see Measures section); (2) covariances between the task-general factors and the task-specific factors within each task. PS = Processing Speed Task; AS = Attention Shifting Task; MR = Mental Rotation Task; Inh = Inhibition Task; $v$ = Drift rate; $a$ = Boundary separation; $t0$ = Non-decision time.

for using the *lavaan.survey* package (Oberski, 2014). Finally, the sociodemographic covariates that were included in the MNLFA scores (see Measures section above) were controlled for in the SEM. Goodness-of-fit was assessed using the root mean square error of approximation (RMSEA) and the comparative fit index (CFI). Following Hu & Bentler (1999), CFI values > .90 and RMSEA values < .08 were interpreted as acceptable model fit and CFI values > .95 and RMSEA values ≤ .06 as good model fit.

**Step 3: Model validation in test set.** After optimizing the model based on the training set, we refit it to the test data. Model fit was assessed the same way as at Step 2. The regression coefficients of these models were interpreted to address our research questions. We controlled for multiple testing in the regression paths based on the false discovery rate (Benjamini & Hochberg, 1995; Cribbie, 2007). We did so separately for tests involving drift rates, non-decision times, and boundary separations, as we had different hypotheses for each of these parameters. In addition, we were interested in determining if standardized effects that fell between -.10 and .10 were consistent with an actual null effect. For regression coefficients falling within these bounds, we therefore used two one-sided tests (TOST) equivalence testing using -.10 and .10 as bounds.

## 2.3 Results

### Model fit

#### DDM

Based on an assessment of model fit, we selected the following good-fitting DDM models for the substantive analysis: 1) *Mental Rotation Task*, the standard model; 2) *Inhibition Task*, the standard model with one set of parameter estimates across conditions; 3) *Attention Shifting Task*, the standard model with one set of parameter estimates across conditions; 4) *Processing Speed Task*, the standard model, but with RTs < 1 s excluded to solve issues with fast outliers. See the supplemental materials for a full overview of the DDM fitting results.

The preregistered simulation-based model fit analysis yielded four (out of 16) correlations between observed and simulated RTs/accuracy that fell below the .80 cut-off: accuracies for Inhibition (.79), Attention Shifting (.73), Processing Speed (.65), and the 75th percentile of RTs for Mental Rotation (.76). However, further analyses showed that all correlations were > .80 when we simulated 100 trials for each task, instead of the same number of trials as the real data. This suggested that the low correlations did not indicate bad parameter recovery, but rather a limitation in the preregistered procedure. Therefore, we decided against further changes to the models or the removal of data points. We provide more details about the model fit procedure, as well as the nature and reason of the deviation, in the supplemental materials (as well as the model fit results for the preregistered and updated approach).

Table 2.2 shows bivariate correlations between DDM parameters and adversity measures. Both material deprivation and household threat showed small, negative associations with drift rates across all four tasks, suggesting that participants with more adversity exposure processed information more slowly. In addition, both material deprivation and household threat were positively associated with boundary separation (indicating more response caution) in all tasks except Mental Rotation, although most of these correlations were very small. Finally, material deprivation and household threat showed a small, negative correlation with non-decision times on the Mental Rotation Task, but not with non-decision times on the other tasks.

**Table 2.2.** Bivariate correlations between DDM parameters and measures of adversity.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Drift Rate** | | | | | | | | | | | | | | |
| 1. Flanker | - | | | | | | | | | | | | | |
| 2. Att. Shift. | 0.43 | - | | | | | | | | | | | | |
| 3. Men. Rot. | 0.27 | 0.30 | - | | | | | | | | | | | |
| 4. Proc. Speed | 0.31 | 0.39 | 0.19 | - | | | | | | | | | | |
| **Boundary Separation** | | | | | | | | | | | | | | |
| 5. Flanker | -0.28 | -0.12 | -0.08 | -0.12 | - | | | | | | | | | |
| 6. Att. Shift. | -0.40 | -0.36 | -0.14 | -0.26 | 0.47 | - | | | | | | | | |
| 7. Men. Rot. | 0.00 | 0.01 | 0.29 | -0.05 | 0.06 | 0.09 | - | | | | | | | |
| 8. Proc. Speed | -0.29 | -0.23 | -0.11 | -0.28 | 0.33 | 0.42 | 0.11 | - | | | | | | |
| **Non-Decision Time** | | | | | | | | | | | | | | |
| 9. Flanker | -0.03 | 0.05 | -0.00 | 0.02 | 0.52 | 0.33 | 0.05 | 0.23 | - | | | | | |
| 10. Att. Shift. | -0.07 | 0.02 | -0.05 | -0.02 | 0.34 | 0.21 | 0.03 | 0.20 | 0.40 | - | | | | |
| 11. Men. Rot. | 0.15 | 0.21 | 0.27 | 0.12 | 0.08 | 0.01 | 0.14 | 0.04 | 0.19 | 0.16 | - | | | |
| 12. Proc. Speed | -0.07 | -0.02 | -0.07 | 0.01 | 0.26 | 0.20 | 0.02 | 0.12 | 0.28 | 0.30 | 0.16 | - | | |
| **Adversity** | | | | | | | | | | | | | | |
| 13. Mat. Dep. | -0.19 | -0.23 | -0.21 | -0.11 | 0.06 | 0.14 | -0.08 | 0.11 | -0.00 | 0.00 | -0.14 | -0.02 | - | |
| 14. Househ. Thr. | -0.12 | -0.15 | -0.10 | -0.10 | 0.02 | 0.06 | -0.03 | 0.07 | -0.03 | -0.02 | -0.08 | -0.02 | 0.26 | - |
| Mean | 2.91 | 1.49 | 0.25 | 1.47 | 2.95 | 2.12 | 2.88 | 2.89 | 0.34 | 0.33 | 1.15 | 1.22 | 0.05 | -0.06 |
| SD | 0.87 | 0.39 | 0.26 | 0.38 | 0.41 | 0.45 | 0.44 | 0.47 | 0.08 | 0.08 | 0.28 | 0.14 | 1.05 | 0.83 |
| Skew | -0.25 | -0.21 | 0.58 | 0.18 | -0.10 | 0.25 | -0.49 | -0.15 | 0.06 | 0.44 | -0.30 | -0.06 | 0.73 | 0.52 |
| Kurtosis | -0.28 | 0.05 | 0.02 | -0.18 | 0.53 | -0.13 | 0.28 | -0.36 | -0.41 | -0.21 | 0.35 | 0.03 | 0.11 | -0.57 |

*Note:* Att. Shift. = Attention Shifting; Men. Rot. = Mental Rotation; Proc. Speed = Processing Speed; Mat. Dep. = Material Deprivation; Househ. Thr. = Household Threat

**SEM**

The SEM model was incrementally constructed in the training data in order to detect any parts that might need adjustment. All parts of the model provided an acceptable to good account of the training data (full training model: CFI = .98, RMSEA = .04). Therefore, we did not make any adjustments to the model before applying it to the test data (N = 9063). The full model also provided a good account of the test data (CFI = .98, RMSEA = .05).

Figure 2.4 presents a simplified overview of the measurement part of the final model in the test data (excluding task-specific covariances and regression paths involving the adversity measures). The factor loadings of the Mental Rotation Task were low for all DDM parameters, suggesting that performance on this task differs substantially from performance on the other tasks. All tasks showed a statistically significant portion of task-specific variance after accounting for task-general effects. Task-general drift rate and task-general boundary separation were negatively correlated ($r = -0.57$), while task-general boundary separation and task-general non-decision time were positively correlated ($r = .71$). These findings show that youth who processed information faster were less cautious in decision-making than those who processed information more slowly, and that more cautious youth were slower in executing non-decision processes (e.g., encoding, response execution) than less cautious youth. Task-specific correlations between DDM parameters of the same tasks ranged between $r = .02$ and $r = .34$.

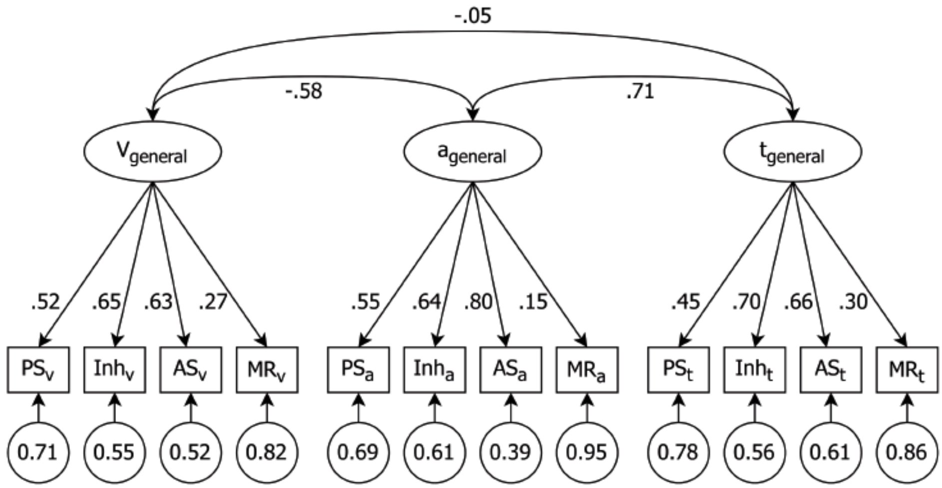



**Figure 2.4.** Simplified overview of the measurement part of the final SEM model, including standardized factor loadings, unstandardized residual variances, and correlations between the general latent factors. Excluding task-specific residual covariances and regression paths (see Figure 2.5). The elipses represent latent task-general factors. The circles represent latent task-specific factors. *v* = drift rate; *a* = boundary separation; *t0* = non-decision time; PS = Processing Speed Task; AS = Attention Shifting Task; MR = Mental Rotation Task; Inh = Inhibition Task.

### Primary analysis

Our primary analysis examined to what extent household threat and material deprivation were associated with task-specific and task-general aspects of speed of information processing (drift rates), response caution (boundary-separations), and task preparation/execution (non-decision times). Task-general effects capture variance shared across tasks, whereas task-specific effects capture variance unique to specific tasks. The results are summarized in Figure 2.5.

For household threat, we found a significant negative association with task-general drift rate ($\beta$ = −0.12, 95% CI = [−0.16, −0.08], $p$ <.001), indicating that participants with more exposure to household threat processed information more slowly in general. All task-specific drift rates were practically equivalent at different levels of household threat. We also found a significant positive association between household threat and task-general boundary separation ($\beta$ = 0.08, 95% CI = [0.04, 0.12], $p$ <.001), indicating that participants with more exposure to household threat generally responded with more caution. In contrast, we found a negative association between household threat and task-specific boundary separation in the Attention Shifting Task ($\beta$ = −0.07, 95% CI = [−0.11, −0.02], $p$ = .013), indicating that participants with more exposure to household threat responded with less caution in this task. The association between house-

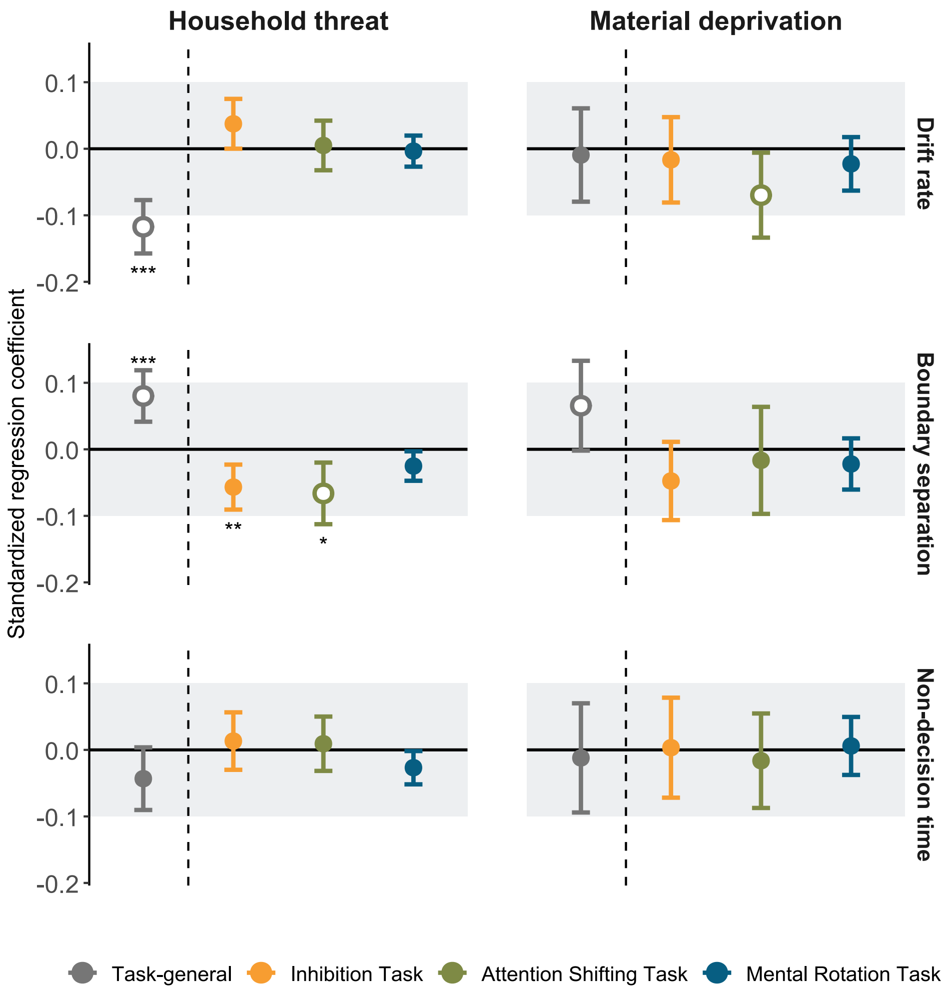



**Figure 2.5.** Results of the structural part of the SEM model testing the effect of household threat and material deprivation on task-specific and task-general DDM parameters. The top row plots the drift rates, the middle row plots the boundary separations, and the bottom row plots the non-decision times. The gray area reflects the area of practical equivalence. Hollow points indicate effects outside the area of practical equivalence. Solid points indicate effects inside the area of practical equivalence. Standard-errors represent 95% confidence intervals. Statistical significance (tested against zero) is indicated with significance asterisks.
* $p < .05$, ** $p < .01$, *** $p < .001$

hold threat and task-specific boundary separation on the Inhibition Task was also significant, but fell in the region of practical equivalence. Both task-general non-decision time and task-specific non-decision times were practically equivalent at different levels of household threat.

For material deprivation, the associations with task-general drift rate, as well as with all task-specific drift rates, were not significantly different from zero. We found evidence for practical equivalence for task-general drift rate and the task-specific drift rates of the Inhibition Task and the Mental Rotation Task. However, we did not find evidence for practical equivalence for the task-specific drift rate of Attention Shifting, suggesting that participants with higher levels of material deprivation might be somewhat slower at shifting attention. The association between material deprivation and task-general boundary separation was neither significantly different from zero ($\beta$ = 0.07, 95% CI = [−0.00, 0.13], $p$ = .091), nor practically equivalent ($p$ = .159). Thus, participants with more exposure to material deprivation might generally respond with somewhat more caution, but the effect size of this relationship is likely not meaningful. All of the task-specific boundary separations were practically equivalent at different levels of material deprivation. Both task-general non-decision time and task-specific non-decision times were practically equivalent at different levels of material deprivation.

### Exploratory analysis

To situate our primary analysis in the context of the broader literature based on raw performance measures, we decided to run a similar SEM model based on raw performance measures of the four cognitive tasks. We used the measures as provided in the ABCD database (Luciana et al., 2018). For the Processing Speed Task, the traditional raw measure is the number of correctly completed trials. For the Mental Rotation Task, the traditional raw measure is the percentage correct divided by the mean response time on correct trials. For the Attention Shifting and Inhibition Task, the traditional raw measure is a composite of accuracy and RT (Slotkin et al., 2012). The model was the same as the primary analysis, with the exception that it included only one task-general factor. Like the primary models, the exploratory model provided a good account of the test data (CFI = 1, RMSEA = 0.04).

The results are summarized in Figure 2.6. Similarly to the primary analysis, household threat was significantly negatively associated with task-general performance. In addition, we found a significant—but practically equivalent—positive association between household threat and task-specific Flanker performance. All of the other effects were practically equivalent at different levels of adversity.

## 2.4 Discussion

Our aim was to better understand how two types of adversity—household threat and material deprivation—are associated with performance differences on three tasks covering inhibition, attention shifting, and mental rotation. First, we used DDM to distinguish between three potential sources for performance differences: 1) the speed of information processing (drift rates), 2) response caution (boundary separation), and 3) the speed of encoding and response execution (non-decision time). Second, we used SEM to investigate if observed differences in each DDM parameter were task-general

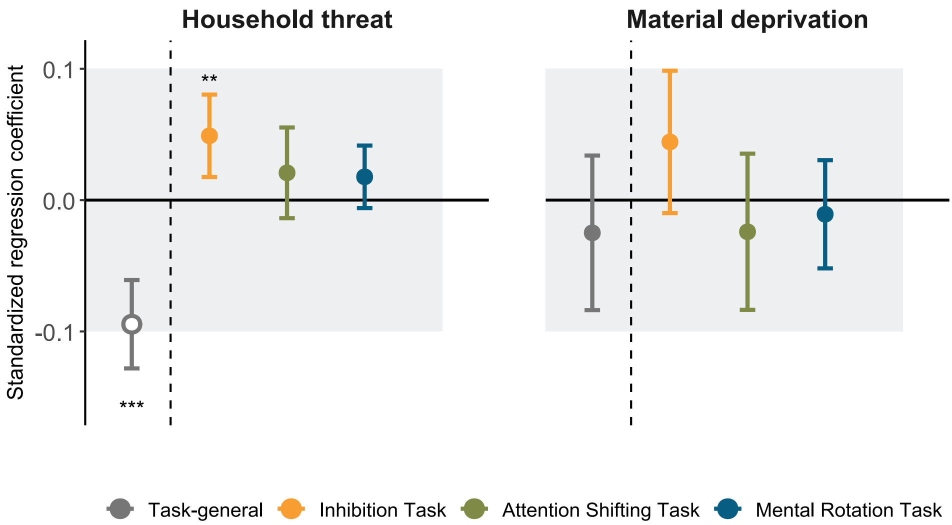



**Figure 2.6.** Exploratory analysis testing the association between household threat and material deprivation on task-specific and task-general raw performance measures. The gray area reflects the area of practical equivalence. Hollow points indicate effects outside the area of practical equivalence. Solid points indicate effects inside the area of practical equivalence. Standard-errors represent 95% confidence intervals. Statistical significance (tested against zero) is indicated with significance stars.
*** $p < .001$, ** $p < .01$, * $p < .05$

(i.e., shared across all tasks) or task-specific (i.e., unique to a specific task). Negative associations between adversity and either task-general or task-specific drift rates would be consistent with existing deficit frameworks. Positive associations between adversity and task-specific drift rates would be consistent with existing adaptation frameworks. In contrast, associations with other DDM parameters, or equivalent drift rates, would not be consistent with either framework.

### Primary findings

Our results provided some support for deficit frameworks, but not for adaptation frameworks. Higher levels of household threat (but not material deprivation) were associated with lower task-general speed of information processing. This was consistent with deficit frameworks, although based on previous literature, we actually expected stronger deficit patterns for deprivation than for threat (Salhi et al., 2021; Sheridan et al., 2020; Sheridan & McLaughlin, 2014). Inconsistent with either deficit or adaptation frameworks, task-specific inhibition and mental rotation abilities were intact. The only exception was the negative association between material deprivation and attention shifting, where we did not find evidence for a significant attention shifting difference, nor for truly intact shifting. Finally, both household threat and material deprivation led to more response caution, although the evidence for material deprivation was weak (not significantly different from zero, but also not practically equivalent to zero). We did not find any differences in task-general or task-specific aspects of task preparation and response execution.

The finding that most task-specific abilities—after accounting for task-general processing speed—were not affected by either type of adversity was striking in light of the existing literature. It suggests that specific executive functions (i.e., inhibition, attention shifting, mental rotation) of youth with more adversity exposure were comparable with those of youth from low-adversity contexts. This is inconsistent with previous interpretations of adversity-related performance differences based on raw performance measures. For example, a previous study showed enhanced attention-shifting performance in youth with more exposure to threat (Young et al., 2022; for similar findings with environmental and caregiver unpredictability, see Fields et al., 2021; Mittal et al., 2015). In addition, youth from adversity have previously been found to perform worse on inhibition tasks (Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015; Noble et al., 2005), and previous investigations in the ABCD study found negative associations between SES and mental rotation (Assari, 2020; Bignardi et al., 2024).

Instead, higher levels of household threat were associated with a lower task-general drift rate. We argue that this is likely to reflect a slower basic speed of processing for three reasons. First, previous studies showed that performance on executive functioning tasks involves basic processing speed (Frischkorn et al., 2019), with one study suggesting that it may be the predominant factor explaining individual differences on executive functioning tasks (Löffler et al., 2024). Second, we included a simple Processing Speed Task to inform and scale each task-general factor. Third, the drift rates of the Flanker and Attention Shifting Task were collapsed across incongruent (switch) and congruent (repeat) trials. Thus, it is likely that the task-general drift rate accounted not only for variance related to incongruent (shift) trials, but also for variance related to the congruent (repeat) trials, which are generally thought to involve mostly basic processing. While we consider the basic processing speed interpretation most likely given these reasons, we note that others have proposed that shared variance among executive functioning tasks predominantly reflects executive attention, or the ability to avoid distraction and to focus and maintain attention (Mashburn et al., 2023; Zelazo & Carlson, 2023). More research is warranted to test these two hypotheses against each other.

Our results align to some extent with two recent investigations. First, Bignardi et al. (2024) conducted a study in three large datasets—among which the ABCD study—in which they used SEM to separate task-general variance from task-specific variance. They found that SES was positively associated with lower task-general performance in all datasets, but after accounting for task-general performance, found many instances of practically equivalent performance. Interestingly, they found negative associations (meaning better performance) between SES and the Flanker and Attention Shifting Task in the ABCD data. Second, Young et al. (2024) examined associations between SES and unpredictability with performance on an achievement task battery, comparing specific subtasks to overall performance across tasks. Similar to our findings, lower SES was associated with lower overall performance, but with intact (or even enhanced) performance on most specific subtasks, relative to the overall effect. However, these studies did not separate cognitive abilities from other processes such as response caution.

Household threat (and to a lesser extent material deprivation) was also associated with more task-general response caution. Traditional assessments could misinterpret this as impaired ability, as it slows down responses. In contrast, task-specific response caution was lower for the Attention Shifting and Inhibition Task (although the latter was practically equivalent). Thus, youth with more exposure to household threat are generally more cautious, but become less cautious specifically when processing conflicting information (i.e., distractions on the Inhibition Task and changing task-demands on the Attention Shifting Task). What might explain these differences? In comparing deficit and adaptation frameworks, we focused mainly on cognitive abilities with a clear performance benchmark (e.g., higher drift rates reflecting better performance). Differences in response caution reflect strategies, not abilities (Frankenhuis, Young, et al., 2020). However, we speculate that these findings could reflect contextually appropriate adaptive responses to threatening conditions. Evidence across multiple species suggests that a high probability of threat tends to increase general response caution (prioritizing accuracy over speed), to avoid costly mistakes (Chittka et al., 2009). However, under acute threat, prioritizing speed over accuracy might be better (e.g., fleeing even though there was no threat). Although the Inhibition and Attention Shifting Task did not signal threat, they did evoke competing demands and conflicting information. In real-life settings, such environmental cues could signal a threat, in which case prioritizing speed over accuracy would facilitate rapid detection and responding (Frankenhuis et al., 2016; Mittal et al., 2015). However, as neither pattern was preregistered, we should calibrate our interpretations accordingly.

### Strengths, limitations, and future directions

The current study has several strengths. First, the analyses were based on the ABCD sample, a large, representative US sample. Second, we developed a framework that can simultaneously account for adversity-related impairments and enhancements and captures cognitive processes that are more theoretically meaningful than raw scores. Third, we used measures of material deprivation and household threat that were corrected for measurement non-invariance using MNLFA, resulting in unbiased estimates of both dimensions of adversity.

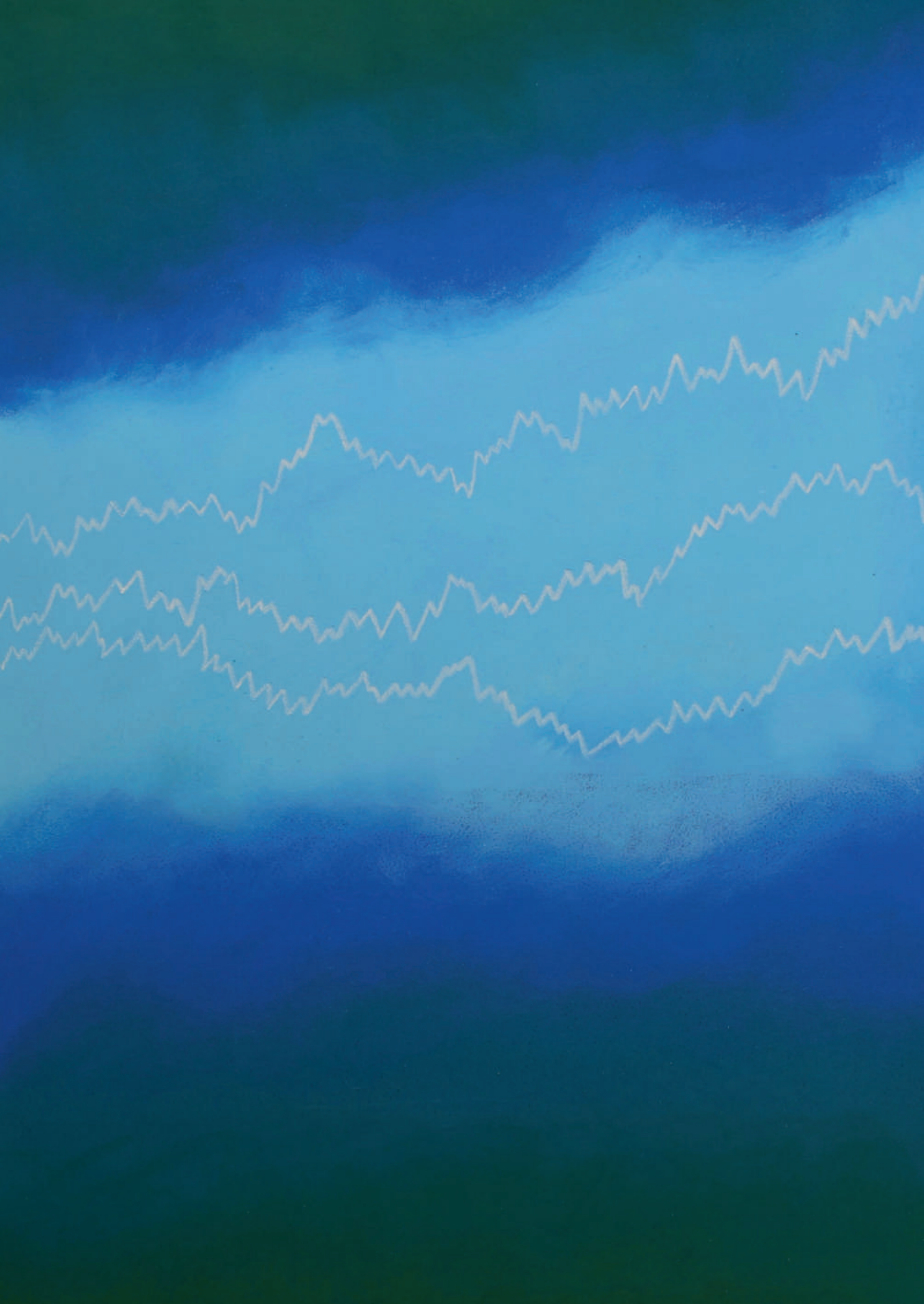
The current study also had limitations. First, we were only able to include three cognitive abilities (aside from processing speed) that were compatible with DDM assumptions. This inevitably excluded many important abilities, which limited the scope of what is captured both in task-general and task-specific processes. Second, because of the low number of trials per task we were unable to separately model the task conditions of the Flanker and Attention Shifting Task. This may have made the task-specific estimates less precise measures of inhibition and attention shifting. Third, despite the enhanced individual variation gained from the MNLFA scores, items composing those scores of household threat and material deprivation were binary, asking for the presence or absence of certain exposures over the last 12 months. Therefore, we were not able to account for the role of frequency and severity of those experiences in that window (let alone over the whole of ontogeny). Fourth, while household threat was child-reported, material deprivation was parent-reported. Thus, the measure of material deprivation might not have fully captured youths' own subjective perception, which may partly explain why household threat was more strongly related to cognitive performance than material deprivation.

Future research can build on this study in a couple of ways. First, it will be important to better understand the processes making up task-general drift rate. To this end, future research should include measures of candidate processes (e.g., basic processing speed, attention maintenance), ideally several measures per process to obtain good latent estimates. In addition, neuro imaging data could be linked directly to DDM parameters to investigate which brain networks are associated with differences in task-general drift rates (e.g., Schubert & Frischkorn, 2020). Second, future research could aim to better understand task-general and task-specific differences in response caution. For example, do youth from adversity show more task-general response caution due to performance anxiety? If so, does such anxiety interfere more with their performance on some tasks than others? Can training programs targeting anxiety boost their performance? Third, our approach could be extended to model developmental trajectories of the cognitive processes as a function of adversity.

Our approach of combining DDM and SEM can also enrich perspectives that promote using culturally-sensitive assessments of executive functioning that relate better to youths pre-existing goals, values, and lived experiences (Doebel, 2020; Miller-Cotto et al., 2022; Niebaum & Munakata, 2023; Nketia et al., 2024; Zuilkowski et al., 2016; also see Zelazo & Carlson, 2023). We agree that more ecologically relevant assessments are needed, but, to the extent that they also rely on response times and accuracy, will suffer from some of the same methodological limitations as traditional tasks. This is exemplified by recent attempts to make task-content more ecologically relevant. While promising, the effects are sometimes difficult to interpret, with different types of content affecting performance in unexpected and inconsistent ways—in some cases helping and in others hindering performance. For instance, testing materials involving money can help to close achievement gaps on working memory tasks (Young et al., 2022), but at the same time harm performance on mathematics exams (Duquennois, 2022; Muskens et al., 2019). This could mean that 1) the effect of these materials on performance is task or domain-specific, and 2) that specific manipulations can have different—even opposing—effects depending on the relevant process. Our approach offers a crucial tool to systematically unpack these differences and to understand how interventions can be best tailored to a child's unique circumstances given a particular cognitive domain.

## 2.5 Conclusion

Taken together, we find that adversity is mostly associated with task-general processes, as well as ability-irrelevant response caution, yet that task-specific abilities are mostly intact. This suggests that traditional cognitive assessments may overestimate the effect of adversity on youth's specific abilities (both impairments and enhancements). Our analytical approach provides a solution. By combining DDM and SEM approaches, we can start to develop a more nuanced understanding of how adversity affects different aspects of cognitive performance among youth and across development. This approach requires large datasets containing multiple cognitive tasks, a requirement that is increasingly feasible with the availability of large, secondary datasets in developmental science (Kievit et al., 2022). Thus, we can develop a more balanced, well-rounded understanding of how adversity shapes cognitive development that integrates both deficit and adaptation perspectives.



# Chapter 3

## Adversity is associated with lower general processing speed rather than specific executive functioning abilities

**This chapter is based on**
Vermeent, S., Schubert, A.-L., & Frankenhuis, W.E. Adversity is associated with lower general processing speed rather than specific executive functioning abilities. *Submitted for publication*. https://doi.org/10.31234/osf.io/kqhf7

## 3.0 Abstract

Exposure to adversity may impair executive functioning (i.e., deficit frameworks), but could also enhance, or leave intact, specific EF abilities (i.e., adaptation frameworks). Both frameworks often use raw performance (e.g., speed) to estimate EF ability. However, this approach (1) conflates different cognitive processes, and (2) generally does not distinguish specific EF abilities from processes that are shared across EF tasks, such as general processing speed. Here, we integrate deficit and adaptation frameworks by building bridges with mathematical and cognitive psychology. Specifically, we use cognitive modeling (Drift Diffusion Modeling) to isolate different cognitive processes: speed of information accumulation, response caution, and speed of stimulus encoding and response execution. We then use structural equation modeling to investigate whether associations between adversity and cognitive processes are task-general or ability-specific. We recruited 1061 participants from the Dutch LISS panel. Participants completed a basic processing speed task, two inhibition tasks, and three attention-shifting tasks. We measured exposure to threat and material deprivation in childhood and adulthood. Exposure to threat (but not material deprivation) in adulthood was negatively associated with task-general processing speed. After accounting for task-general processes, remaining variance was not related to either inhibition or attention-shifting ability. Non-preregistered analyses showed that childhood exposure to material deprivation and threat were negatively associated with (1) general processing speed, and (2) task-specific information accumulation. The latter reflected unique features of individual tasks, rather than specific EF abilities. Taken together, these results suggest that adversity researchers overestimate associations between adversity and specific EF abilities when analyzing raw performance.



**Author contributions**

All authors were involved in conceptualizing the study. SV accessed and analyzed the data, and wrote the first draft of the manuscript. All authors provided feedback on the manuscript.

## 3.1 Introduction

Psychologists have used two main frameworks to explain associations between adversity exposure—such as material deprivation or threat—and executive functioning (EF); one focuses on deficits, the other on adaptation. Deficit frameworks predict that adverse experiences impair brain structure and function in ways that undermine social and cognitive abilities (Farah et al., 2006; Merz et al., 2019; Rosen et al., 2018; Ursache & Noble, 2016a). Adaptation frameworks predict that adverse experiences can lead to the development of intact or enhanced EF abilities, specifically those abilities useful for solving challenges posed by adverse environments (Ellis et al., 2022; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013).

Although the two frameworks make opposing predictions in some cases, they share the goal to reveal how adversity influences specific EF abilities (Ellis et al., 2022; Farah et al., 2006; Frankenhuis, Young, et al., 2020; D. Johnson et al., 2021). Knowledge about impaired abilities provides valuable targets for interventions designed to bridge achievement gaps. Conversely, knowledge about enhanced abilities can be leveraged in school or work contexts, for instance, by redesigning learning settings to match people's unique strengths.

Research within both frameworks typically employs the same inferential strategy, by estimating specific EF abilities based on raw performance on EF tasks (e.g., response speed or accuracy). For instance, on the one hand, children who live in less favorable environments tend to be slower and less accurate on inhibition tasks—which has been interpreted as an impaired ability to inhibit distractions (Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015; Noble et al., 2005). On the other hand, some studies report that adolescents and young adults with more exposure to threat and unpredictability might perform better at shifting their attention between different tasks, without sacrificing much performance on each task—which may be a useful adaptation for tracking information in fast changing and dangerous environments (Fields et al., 2021; Howard et al., 2020; Mittal et al., 2015; Nweze et al., 2021; Young et al., 2022). In this paper, we will challenge the inferential strategy of estimating EF abilities based on raw performance. We advocate for cognitive modeling as a viable alternative approach to estimating EF abilities.

#### Methodological challenges for measuring specific EF abilities

Recent research shows that raw performance on EF tasks does not necessarily capture specific EF abilities, for two related reasons. First, raw performance on EF tasks is also influenced by processes other than the ability of interest, such as response caution and the speed of response execution (Hedge et al., 2022; Ratcliff & McKoon, 2008; Stahl et al., 2014). Thus, slower (or faster) responses in people with more adversity exposure need not reflect EF ability, but could, for instance, reflect a tendency to respond more cautiously (Vermeent et al., 2024). Second, raw performance on EF tasks is also influenced by general processes, such as basic speed of information processing (Frischkorn et al., 2019; Löffler et al., 2024; Weigard et al., 2021). Thus, lower performance on multiple EF tasks may reflect a difference in a general process, rather than specific EF abilities (Bignardi et al., 2024; Vermeent et al., 2024; Young et al., 2024).

The standard inferential strategy of estimating EF abilities from raw performance limits the effectiveness of interventions. For example, some interventions include training modules for specific EF tasks if performance on a task is below some prespecified cut-off (Distefano et al., 2021). Performance on inhibition tasks in particular is thought to mediate the association between adversity exposure and achievement (Taylor & Barch, 2022). At face value, this would suggest interventions targeting inhibition ability (e.g., inhibition training, which might involve removing distractions from the environment) could positively impact achievement outcomes. However, if impaired inhibition ability is *not* the root cause of performance differences, but rather general processes are (such as the speed of information processing), the impact of such interventions will likely be limited. Therefore, adversity research should move beyond the use of raw performance to better understand the associations between adversity exposures and specific EF abilities.

### Separating EF ability from other decision-making processes

Raw performance confounds multiple stages of processing: from initial preparations (e.g., stimulus encoding), to processing task-relevant information, to deciding on and executing a response (Forstmann et al., 2016; Ratcliff, 1978; Wagenmakers, 2009). Individual differences in response times and error rates also depend on how cautiously people make decisions. More cautious people tend to respond more slowly to increase their accuracy, which does not necessarily imply lower ability (Voss et al., 2004). Thus, understanding associations between adversity and specific EF abilities requires isolating ability-relevant processing (e.g., inhibition, attention-shifting) from ability-irrelevant processes that contribute to response times.

Cognitive modeling can distinguish between abilities and ability-irrelevant processes. Here, we focus on the Drift Diffusion Model (DDM; Forstmann et al., 2016; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Wagenmakers, 2009), which can be applied to many widely used EF tasks. The DDM assumes that people accumulate evidence for one of two responses (e.g., left or right button) until they have acquired sufficient evidence to make a decision (see Figure 3.1). When applied to trial-level response times and accuracy data, the DDM estimates four parameters that represent distinct cognitive processes (Voss et al., 2004). The *drift rate* reflects the rate at which people accumulate evidence for the correct response; thus, it measures the efficiency of information processing. A higher drift rate leads to faster responses and higher accuracy. When applied to EF tasks, the drift rate captures individual differences in specific EF abilities (although drift rate may also capture more general processes, see below; Löffler et al., 2024; Vermeent et al., 2024; Weigard et al., 2021). The *boundary separation* reflects

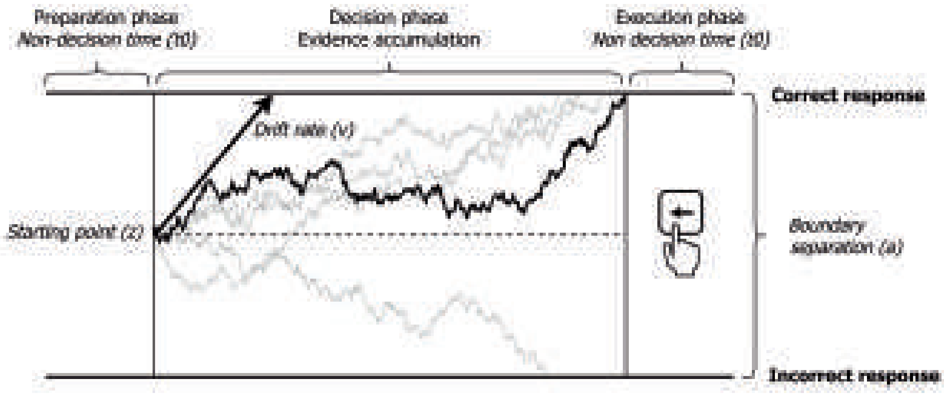



**Figure 3.1.** Visual representation of the Drift Diffusion Model (DDM). The DDM assumes that people accumulate evidence for one of two responses until they have acquired sufficient evidence to make a decision. Each squiggly line represents the evidence accumulation process on a single trial. The thick, horizontal lines represent the decision boundaries corresponding to the correct response (upper boundary) and the incorrect response (lower boundary). Applying the DDM to trial-level performance data (response times and accuracies) yields four parameters representing distinct cognitive processes: (1) **drift rate**: Average rate of evidence accumulation across trials; measuring the efficiency of information processing. (2) **boundary separation**: The distance between the two decision boundaries; measuring response caution. (3) **Non-decision time**: A combination of the preparation speed (e.g., stimulus encoding) and response execution speed (e.g., time spent pressing the button); measuring speed of non-decision processes. **starting point**: The starting point of the evidence accumulation process; measuring response bias (not considered here). Figure adopted from Vermeent et al. (2024), licensed under CC BY 4.0.

the width between the two decision boundaries; it measures a person's response caution. A higher boundary separation leads to slower responses and higher accuracy. The *non-decision time* reflects a combination of preparation time (e.g., stimulus encoding) as well as execution time (e.g., time spent pressing the button). A larger non-decision time leads to slower responses without a change in accuracy. Finally, the *starting-point* allows for initial bias for one of two responses (e.g., happy or angry faces). In this paper, we will not consider the starting point, because people cannot have biases towards correct or incorrect responses.

### Separating EF ability from general processing speed

Recent studies show that a single general factor accounts for most of the variance in drift rates across multiple tasks (Lerche et al., 2020; Löffler et al., 2024; Weigard et al., 2021). This general factor appears to be largely stable across time (Schubert et al., 2016; Weigard et al., 2021). It arises also in analyses of tasks thought to measure different abilities (Schmiedek et al., 2007; Schmitz & Wilhelm, 2016). Even if all tasks are designed to measure the same EF ability (Hedge et al., 2022), variance in the general factor is fully explained by general processing speed instead of any specific EF ability

(Löffler et al., 2024). These findings suggest that performance on standard EF tasks largely reflects basic speed of processing, rather than EF abilities.

Recent findings suggest that the effects of adversity are more general than specific. For instance, youth with lower SES and more exposure to adversity show lower task-general raw performance scores across measures of EF, memory, and intelligence, while task-specific performance (after accounting for general variance) is often practically equivalent to zero (Bignardi et al., 2024; Young et al., 2024). In addition, youth with more exposure to household threat showed lower task-general processing speed (based on drift rates), while differences in task-specific drift rates were practically equivalent to zero (Vermeent et al., 2024). These findings suggest that previous research may have overestimated the associations between adversity and specific abilities.

A consequence of both sources of measurement impurity is that estimating specific EF abilities requires latent variable models including two or more measures of the same ability; yet, most studies in adversity research include only one task per EF ability. This makes this work less suitable to investigate whether adversity is associated with specific EF abilities after accounting for general processes. This is important for both deficit-based and adaptation-based interventions, which aim to remediate or leverage impaired or enhanced abilities, respectively. Moreover, while deficit frameworks predict impairments in both specific abilities and general processes (Vermeent et al., 2024), adaptation frameworks predict that specific types of adversity lead to adaptations in specific EF abilities. For example, it has been proposed that unpredictable and dangerous environments may positively affect the ability to rapidly shift attention, promoting the detection of sudden threats and seizing fleeting opportunities (Frankenhuis, Young, et al., 2020; Mittal et al., 2015). In the current research, therefore, we investigate the associations between adversity and two latent EF abilities—–inhibition and attention shifting—–after accounting for task-general processing speed. We will use multiple tasks for each ability to obtain better estimates of specific abilities.

**The current study**

Using a combination of DDM and structural equation modeling, we address three central questions. First, what is the association of adversity exposure in adulthood with general processing speed that is contributing to performance across all cognitive tasks? Second, what is the association of adversity exposure in adulthood with inhibition and attention-shifting abilities after accounting for general processing speed? Third, what is the association of adversity exposure in adulthood with general and/or EF-specific response caution? We focus on two types of adversity: material deprivation (a lack of access to material resources) and threat (the potential for harm imposed by others). In previous research, both types of adversity were associated with performance on tasks used to measure EF (e.g., Fields et al., 2021; Schäfer et al., 2022; Sheridan et al., 2022;

Vermeent et al., 2024; Young et al., 2022). In addition, both are central to recent dimensional models of adversity (McLaughlin et al., 2021; Sheridan & McLaughlin, 2014).

We preregistered predictions of deficit and adaptation frameworks. Deficit frameworks predict negative associations between adversity and both the general speed of processing as well as specific abilities. Adaptation frameworks predict intact or enhanced attention shifting ability, but not intact or enhanced inhibition ability. We only interpret intact attention shifting ability as an adaptation if inhibition ability is also lower. Intact performance on both abilities, however, would not provide sufficient evidence of adaptation, as it might instead suggest no association between adversity exposure and these abilities. In line with previous work (Vermeent et al., 2024), we further predict that threat exposure is associated with higher task-general response caution.

## 3.2 Methods

### Participants

We tested a total of 1061 participants from the Dutch Longitudinal Internet Studies for the Social Sciences (LISS) panel (Scherpenzeel, 2011). The LISS panel is an invitation-only, representative sample of the Dutch population consisting of approximately 7,500 individuals across 5,000 households. LISS participants complete a yearly core battery of questionnaires about various domains of life, including one's financial situation over the past year. In addition, LISS participants have the option to participate in further monthly studies on different topics. We estimated power based on Kretzschmar & Gignac (2019). With $\alpha = 0.05$ and assuming moderate reliability of our cognitive measures, we would have > 90% power to detect small effect sizes ($\beta = 0.1$) with a sample size ranging between N = 730 and N = 980. Therefore, we aimed for a total sample size of N = 1,000.

The current study took place between May and August 2024. Participants were able to complete the cognitive tasks during two or more sessions to increase participation rates. People were eligible for the study if they were between 18 and 55 years old, and had agreed to linking their LISS data to government microdata (not relevant here, but for a different study). We excluded participants who did not have data on any cognitive task. The final sample after exclusions consisted of 1056 participants (see Table 3.1).

**Table 3.1.** Descriptive statistics for the final sample.

| Category | Statistic |
|---|---|
| Mean age (SD) | 39.3 (11.1) |
| Sex (% Female) | 55.7 |
| Highest completed education (%) | |
| primary school | 3 |
| vmbo (intermediate secondary education) | 11.9 |
| havo/vwo (higher secondary education) | 15.4 |
| mbo (intermediate vocational education) | 39.1 |
| hbo (higher vocational education) | 36.9 |
| wo (university) | 27.9 |
| other | 1.6 |
| missing | 3.3 |
| Mean number of waves (SD) | |
| Material deprivation | 8.7 (4.9) |
| Threat | 3.5 (2.3) |

**Adversity measures**

We preregistered our approach for computing adversity composite scores based on observed correlations between measures. If all measures of an adversity type (i.e., material deprivation and threat) correlated .60 or higher (indicating "strong" correlations, as was the case for material deprivation in adulthood), we calculated a uniformly weighted average. If one or more correlations were lower than .60 (as was the case for neighborhood threat in adulthood), we applied Principal Component Analysis to the separate measures and extracted only the first principal component score. See section 1 of the supplemental materials for frequency distributions of all adversity measures.

**Neighborhood threat in adulthood**

Following our preregistration, we measured exposure to neighborhood violence in adulthood using two measures of perceived neighborhood crime, and one measure of crime victimization. Participants completed the Neighborhood Violence Scale (Frankenhuis, Young, et al., 2020; Frankenhuis & Bijlstra, 2018). This scale includes seven items about current perceived neighborhood threat (e.g., "Where I live, it is important to be able to defend yourself against physical harm"), on a scale of 1 ("Completely disagree") to 7 ("Completely agree"). These items were averaged and standardized. We also included four items on perceived neighborhood threat from the LISS archive (six waves: https://doi.org/10.17026/dans-zch-j8xt). Participants reported how frequently they: (1) "avoid certain areas in your place of residence because you perceive them as unsafe", (2) "do not respond to a call at the door because you feel that it is unsafe", (3) "leave valuable items at home to avoid theft or robbery in the street?", (4) "make a detour, by car or on foot, to avoid unsafe areas?". The scale of these four items ranged from 1 ("(Almost) never"), 2 ("Sometimes"), to 3 ("Often"). We intended to include these items again in the current study, but due to an oversight this did not happen. This meant that we did not have data on these items for participants who never participated in any of the six waves described above (missing N = 348). We decided to conduct all analyses involving threat based only on this subset of the data (N = 708), because *a priori* power remained above 80% (Kretzschmar & Gignac, 2019). We summed the items within each wave, and then calculated an average across waves for which participants had data.

We computed crime victimization in adulthood based on seven items from the LISS archive (six waves: https://doi.org/10.17026/dans-zch-j8xt) in which participants reported whether or not they had been the victim of seven types of crime in the last two years: (1) burglary or attempted burglary; (2) theft from their car; (3) theft of their wallet or purse, handbag, or other personal possession; (4) wreckage of their car or other private property; (5) intimidation by any other means; (6) maltreatment of such serious nature that it required medical attention; and (7) maltreatment that did not require medical attention. We also included these items again in the current study, with people answering them about the last two years. We computed the total number of distinct crimes that participants were exposed to at any moment in time (a 'variety score'; Sweeten, 2012).

The correlations between the two measures of perceived neighborhood crime and crime victimization were low (see Table 3.2). Therefore, following our preregistration, we used Principal Component Analysis to extract the first principal component score. This score accounted for 23 % of the variance in the three measures.

**Material deprivation in adulthood**

We derived measures of material deprivation in adulthood from the LISS archive, using the yearly recurring core study on household and personal income (16 waves: https://doi.org/10.57990/1gr4-bf42). First, participants reported how difficult it currently is to live off the income of their household, on a scale from 0 (very hard) to 10 (very easy). Second, participants reported which of the following statements best described their current financial situation: (1) "we are accumulating debt"; (2) "we are somewhat eating into savings"; (3) "we are just managing to make ends meet"; (4) "we have a little bit of money to spare"; (5) "we have a lot of money to spare". Third, participants reported which of the following applied to their current financial situation (0 = no, 1 = yes): (1) "having trouble making ends meet"; (2) "unable to quickly replace things that break"; (3) "having to lend money for necessary expenditures"; (4) "running behind in paying rent/mortgage or general utilities"; (5) "debt collector/bailiff at the door in the last month"; and (6) "received financial support from family or friends in the last month". We recoded responses so that higher scores indicated more perceived scarcity.

We first reverse-coded and averaged each measure separately across waves, and then scaled them. As all item correlations were > .60 (see Table 3.2), we computed a uniformly weighted average.

**Table 3.2.** Spearman correlations between the main independent variables and covariates.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Living off income | - | | | | | | | | | | | |
| 2. Financial troubles | 0.67 | - | | | | | | | | | | |
| 3. Current situation | 0.73 | 0.65 | - | | | | | | | | | |
| 4. Perceived scarcity | 0.96 | 0.73 | 0.79 | - | | | | | | | | |
| 5. Neighborhood safety | 0.20 | 0.17 | 0.16 | 0.17 | - | | | | | | | |
| 6. Neighborhood Violence Scale | 0.25 | 0.19 | 0.23 | 0.19 | 0.24 | - | | | | | | |
| 7. Crime victimization | 0.11 | 0.21 | 0.17 | 0.29 | 0.08 | 0.08 | - | | | | | |
| 8. Threat | 0.29 | 0.24 | 0.24 | 0.27 | 0.71 | 0.79 | 0.30 | - | | | | |
| 9. Childhood adversity | 0.22 | 0.14 | 0.18 | 0.16 | 0.14 | 0.38 | 0.05 | 0.32 | - | | | |
| 10. Sex assigned at birth | 0.08 | 0.03 | 0.02 | 0.04 | 0.37 | 0.03 | -0.01 | 0.21 | -0.05 | - | | |
| 11. Age | -0.05 | -0.01 | 0.03 | 0.07 | -0.04 | 0.03 | 0.25 | 0.01 | 0.19 | -0.02 | - | |
| 12. Highest education | -0.23 | -0.16 | -0.20 | -0.23 | -0.04 | -0.09 | 0.03 | -0.08 | -0.12 | -0.00 | -0.06 | - |
| Mean | 4.34 | 1.32 | 2.36 | 2.45 | 1.59 | 2.50 | 0.86 | -0.00 | -0.00 | 0.56 | 0.00 | -0.00 |
| SD | 1.62 | 0.59 | 0.75 | 0.91 | 1.58 | 0.97 | 1.23 | 0.71 | 1.00 | 0.50 | 1.00 | 1.00 |
| Min | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 0.00 | -1.20 | -1.41 | 0.00 | -2.01 | -2.42 |
| Max | 11.00 | 5.00 | 5.00 | 5.96 | 8.00 | 6.57 | 7.00 | 3.93 | 4.23 | 1.00 | 1.50 | 3.18 |
| Skew | 0.67 | 2.72 | 0.44 | 0.85 | 1.20 | 1.11 | 1.60 | 1.25 | 1.08 | -0.23 | -0.28 | -0.07 |
| Kurtosis | 1.05 | 8.79 | 0.02 | 0.84 | 1.30 | 1.21 | 2.30 | 2.28 | 0.97 | -1.95 | -1.05 | 0.26 |

**Cognitive measures**

We programmed six cognitive tasks in jsPsych 7.3 (De Leeuw, 2015): two inhibition tasks, three attention shifting tasks, and one basic processing speed task. At the start of the session, participants entered fullscreen mode to avoid distractions from other browser tabs. The tasks were presented against a light-gray background. See Section 2 of the supplemental materials for information on condition manipulation checks, split-half reliability estimates, and bivariate correlations between tasks.

*Flanker Task.* This task measures inhibition of distractor interference (B. A. Eriksen & Eriksen, 1974). On each trial, participants saw five arrows side-by-side horizontally, pointing either left or right. Their task was to indicate the direction of the central arrow. The arrows were randomly presented 300 pixels above or below the center of the screen. On 50% of the trials, all arrows pointed in the same direction (congruent trials). On the other half, the arrows surrounding the central arrow pointed in the opposite direction (incongruent trials). Participants first completed eight practice trials, followed by two test blocks of 32 trials each, for a total of 64 trials.

*Simon Task.* This task measures inhibition of prepotent responses (Simon & Wolf, 1963). On each trial, participants saw the word “LEFT” or “RIGHT” (printed in Dutch), presented either on the left or right side of the screen. Their task was to press the ‘A’ key if the word was “LEFT” and the ‘L’ key if the word was “RIGHT”, regardless of the location on the screen. On 50% of the trials, the word matched the location (e.g., the word “LEFT” presented on the left side; congruent trials). On the other half, the word did not match the location (e.g., the word “LEFT” presented on the right side; incongruent trials). Participants first completed eight practice trials, followed by two test blocks of 32 trials each, for a total of 64 trials.

*Color-shape Task.* This task measures the ability to shift attention between different tasks (Miyake et al., 2000). On each trial, participants saw a square or a circle in the center of the screen. This square or circle was either blue or yellow. Depending on the task rule printed above the stimulus, their task was to classify the stimulus based on its shape or color. On 50% of the trials, the rule was the same as on the preceding trial (repeat trials). On the other half, the rule was different than on the preceding trial (switch trials). The same stimulus was never presented more than twice in a row, and there were never more than three repeat or switch trials in a row. Participants first completed eight practice trials, followed by two test blocks of 32 trials each, for a total of 64 trials.

*Animacy-size Task.* This task measures the ability to shift attention between tasks (Arrington & Logan, 2004). On each trial, participants saw a single noun (in Dutch) referring to an animal or object (adopted from Braem, 2017). Depending on the task rule presented on the screen, their task was to classify the noun based on whether it referred to a living or non-living thing (e.g., wasp vs. piano), or whether it referred to something that was smaller or larger than a soccer ball (e.g., ring vs. dolphin). On 50% of the trials, the rule was the same as on the preceding trial (repeat trials). On the other half, the rule was different than on the preceding trial (switch trials). There were never more than three repeat or switch trials in a row. Participants first completed eight practice trials, followed by two test blocks of 32 trials each, for a total of 64 trials.

*Global-local Task.* This task measures the ability to shift attention between tasks. We adapted the stimuli from Huizinga et al. (2010). On each trial, participants saw a large square or rectangle composed of 16 small squares or rectangles. The stimulus was flanked on both side by a drawing of an elephant or mouse, which was presented 1,000 ms prior to the appearance of the stimulus. If the stimulus was flanked by the image of an elephant (50% of trials), participants had to indicate whether the global image was a square or rectangle. If the stimulus was flanked by the image of a mouse (50% of trials), participants had to indicate whether the local images were squares or rectangles. On 50 % of the trials, the rule was the same as on the preceding trial (repeat trials). On the other half, the rule was different than on the preceding trial (switch trials). Finally, the stimuli were congruent on 50% of the trials (e.g., large square consisting of small squares) and incongruent on the other half (e.g., large square consisting of small rectangles). Congruency, task rule (switch vs. repeat), or focus (global vs. local) where never repeated more than three times in a row. Participants first completed eight practice trials, followed by two test blocks of 32 trials each, for a total of 64 trials.

*Posner Task.* This task measures basic speed of processing (Posner & Mitchell, 1967). On each trial, participants saw two letters in the center of the screen, drawn from the set A, B, F, H, Q, a, b, f, h, and q. Their task was to indicate whether the letters were the same (e.g., "AA", "bB") or different (e.g., "AQ", "Fh"). On 50% of the trials, the letters were the same, and on the other half they were different. Participants first completed eight practice trials, followed by two test blocks of 40 trials each, for a total of 80 trials.

### State anxiety

Participants reported their state anxiety after each cognitive task; thus, six times in total. We measured state anxiety using the shortened version of the State-Trait Anxiety Inventory (Bij et al., 2003; Marteau & Bekker, 1992), which asks participants how calm, tense, upset, relaxed, content, and worried they currently feel, on a scale of 1 ("not al all") to 4 ("very much"). We recoded (if necessary) and then summed the answers with higher scores reflecting more state anxiety.

### Environmental noise

Participants reported the level of environmental noise after each cognitive task; thus, six times in total. We measured environmental noise using a single item, rated on a scale of 1 to 5: "How much noise was there in your environment during the game?"

**Confounds**

We used Directed Acyclic Graphs to identify potential confounds of the key estimands, which were the associations between self-reported threat and material deprivation in adulthood with cognitive outcomes. A Directed Acyclic Graph is a visual overview of assumptions about how variables are causally related. They are graphs consisting of nodes (variables) and directed arrows (causal pathways). An arrow between two variables represents the assumption that experimentally manipulating the variable at the origin of the arrow will change the variable at the end of the arrow (but not the other way around). Directed Acyclic Graphs aid in the identification of variables that need to be adjusted for in the statistical models (i.e., confounders with arrows to both the main predictor and the outcome), and, equally importantly, which variables should not be adjusted for (i.e., colliders and mediators). For a detailed explanation, see Rohrer (2018).

The set of potential confounders consisted of (1) age (Salthouse, 2016, 2019; Starns & Ratcliff, 2010); (2) education (Hofmarcher, 2021); (3) sex (Ning et al., 2023); (4) childhood adversity exposure (material deprivation and threat combined) (Bos et al., 2009; Goodman et al., 2019); and (5) potential causal relations between recent material deprivation and recent threat (Bywaters et al., 2016; Lacey et al., 2022; Ning et al., 2023). We made these decisions based on previous literature, or, in cases of doubt, by statistically testing support for specific relations using data from a previous LISS study (Vermeent et al., 2025). Specifically, following our preregistration, we used the *dagitty* R package (Ankan et al., 2021; Textor et al., 2016) to test whether independencies between specific variables implied by the Directed Acyclic Graph (conditional on the other causal pathways) were supported by these data (see the preregistration for more details).

Figure 3.2 depicts the final preregistered Directed Acyclic Graph for material deprivation and threat exposure in adulthood. For threat exposure in adulthood, our statistical model controlled for age, sex, childhood adversity, and recent material deprivation. We included material deprivation in adulthood based on studies showing that material deprivation tends to precede exposure to adversities such as threat (Bywaters et al., 2016; Lacey et al., 2022; Ning et al., 2023). We did not control for education given a lack of theoretical and statistical support for this effect (see preregistration). To explore the effect of our assumption that material deprivation tends to precede threat, we also present a secondary model excluding material deprivation in adulthood as a confound. As preregistered, we will not base our main conclusions on this secondary model.

For material deprivation in adulthood, our primary statistical model controlled for age, education, sex, and childhood adversity exposure based on previous studies (see above). We did not control for threat exposure in adulthood, as, following the Directed Acyclic Graph, we assume that threat exposure mediates the effect of recent material deprivation on cognitive processes. If true, including threat exposure in the model would make the association between material deprivation and cognitive processes an indirect effect, rather than a total effect (which was our estimand). To explore the effect of our assumption that threat mediates the effect of material deprivation on cognitive processes, we also present a secondary model that includes threat in adulthood as a confound. As preregistered, we will not base our main conclusions on this secondary model.

### Procedure

We obtained ethical approval from the Ethics Review Board of the Faculty of Social & Behavioral Sciences of Utrecht University (FETC20-490) and the Ethics commit-

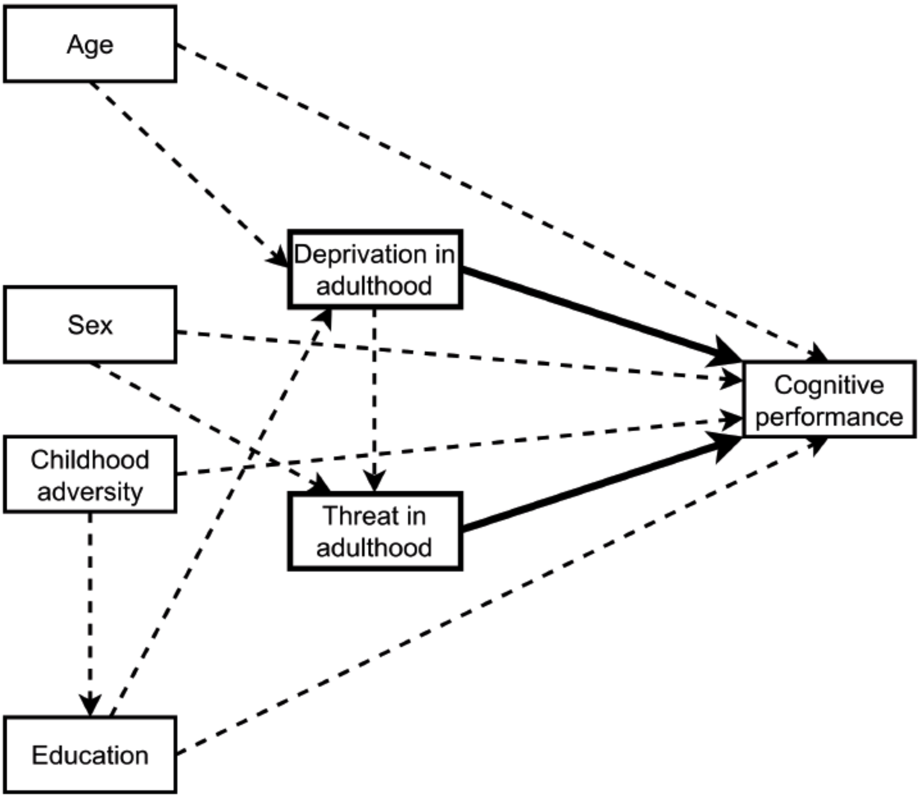



**Figure 3.2.** Direct Acyclic Graphs (DAG) depicting our causal assumptions about the main estimands. the rectangles (nodes) depict variables, and directed arrows depict assumed causal pathways. An arrow between two variables represents the assumption that experimentally manipulating the variable at the origin of the arrow will change the variable at the end of the arrow (but not the other way around). The statistical models control for confounding variables (depicted as dashed paths), which have an arrow both to the independent variable and the dependent variable.

tee for research in the Sciences and Life Sciences of the University of Amsterdam (FNWI-41_2023). The study was implemented on the LISS platform, and participants could only complete the study on a laptop or desktop PC. Participants started with the six cognitive tasks, in randomized order. After each task, they indicated their state anxiety and the level of environmental noise. Then, participants had the option to either continue to the next task or complete the other tasks at a later point in time. 86 % of participants completed all tasks in a single session, and 14 % of participants took two or more sessions. Jointly, the cognitive tasks took around 25 minutes to complete. After finishing all cognitive tasks, participants completed questionnaires about neighborhood threat and material deprivation during childhood and, adulthood exposure to crime victimization in the past two years. They also completed questionnaires on impulsivity, self-control, and future orientation, which are not considered here because they fall outside of the scope of the present study. As some participants had already completed all of these questionnaires in a different data collection (about six months prior to the current study), we only presented the questionnaires to new participants. Finally, all participants answered a few standard LISS evaluation questions about the study, giving them an opportunity to provide written feedback.

## Analysis plan

### Data cleaning

For all tasks, we first removed any response times > 10 seconds. This step was not preregistered but was necessary given that we did not specify a response time-out. Consequently, a small portion of response times lasted up to several minutes, likely reflecting breaks or interruptions (between 0.02 % and 0.25 % for all tasks). We removed these first to prevent them from biasing our preregistered exclusion of outliers. Next, we applied two preregistered exclusion criteria. First, we removed trials with response times < 250 ms and trials with response times more than 3.2 SD above the intra-individual log-transformed mean response time. Second, if participants performed at chance level on a particular task, we excluded the data for that task only. We set the cut-off for chance performance based on accuracy at the 97.5 % tail of the binomial distribution one would obtain if guessing.

### DDM estimation

We used a Hierarchical Bayesian implementation of the DDM (Vandekerckhove et al., 2011; Wiecki et al., 2013) to leverage group-level information for individual parameter estimation. We applied this model to each task separately. For the Posner Task, we estimated a single drift rate, boundary separation, and non-decision time for each participant. For the other five tasks, we estimated drift rates, boundary separation, and non-decision times separately for congruent (repeat) and incongruent (switch) trials.

The DDM models were fit using the *runjags* package (Denwood, 2016). We fit each model with three Markov Chain Monte Carlo chains. We used 2,000 burn-in samples and 10,000 additional samples, retaining every 10th sample, resulting in a total of

3,000 posterior samples. Model convergence and fit was good across all DDM models (see section 3 of the supplemental materials).

We accounted for two potential sources of measurement error in the DDM estimates: (1) the level of environmental noise and (2) within-participant differences from mean state anxiety between the tasks (see above). We used linear regression to residualize the variance of noise and anxiety out of all the drift diffusion estimates. See Section 4 of the supplemental materials for more information.

**Structural equation modeling**

We constructed the full structural equation model sequentially. First, we optimized the fit of the drift rate, boundary separation, and non-decision time sub-models. Second, we combined these three models into a single measurement model. Third, we added the regression paths between measures of adversity in adulthood and the latent factors. To assess goodness-of-fit, we used the root mean square error of approximation (RMSEA) and the comparative fit index (CFI). CFI values > 0.90 (> 0.95) and RMSEA values < 0.08 (≤ 0.06) were interpreted as acceptable (good) fit.

Each DDM parameter sub-model was a bi-factor model including the parameter estimates of all tasks as manifest variables. We estimated a general factor accounting for variance in all manifest variables. A second ability-specific inhibition factor loaded on the parameter estimates of incongruent trials of the Flanker and Simon tasks. A third specific attention shifting factor loaded on the parameter estimates of switch trials of the Color-shape, Global-local, and Animacy-size tasks. These two factors were allowed to covary. We also estimated covariances between the conditions of each task. For each sub-model, we compared the fit of this initial model to a version with a common EF factor loading on incongruent/switch trials of all tasks. We deemed the second model a better fit when we observed a significant chi squared change test and an AIC value difference > 10.

For the drift rate model, we interpreted the general factor as basic speed of processing, and the ability-specific factors as reflecting inhibition and attention-shifting ability (or common EF in the case of the second model). For the boundary separation and non-decision time models, we interpreted the general factor as general response caution/speed of non-decision processes, and the ability-specific factors as reflecting response caution/speed of non-decision processes specific to conflict trials.

After optimizing each sub-model, we combined them into a joint measurement model. We allowed the general latent factors (processing speed, general caution, and general non-decision time) to covary, as well as the ability-specific latent factors (separately for inhibition and attention shifting, unless a common factor was favored).

Finally, we constructed two versions of the final model, one to estimate the association between material deprivation in adulthood and the outcome measures, and one to estimate the association between threat exposure in adulthood and the outcome measures. In both models, we regressed the adversity measure on each latent factor, together with the control variables (see the section on confounds for more details). We report indirect effects of control variables in the supplemental materials (section 5).

#### Inferential criteria

We tested for three types of associations between different types of adversity with EF ability and processing speed (as indexed by drift rates): enhancements, impairments, or practical equivalence. We defined enhancements as a positive association between adversity and drift rates. We defined impairments as a negative association between adversity and drift rate. We defined practical equivalence as standardized regression coefficients that significantly fell between −0.10 and 0.10. We tested this using the Two One-Sided T-tests procedure (Lakens et al., 2018). This test evaluates whether the obtained effect is significantly larger than the lower bound *and* significantly smaller than the upper bound. Thus, it affords conclusions about practical equivalence based on significant *p*-values, rather than (invalidly) inferring the absence of an effect on the basis of non-significant results.

#### Transparency and openness

We have reported how we determined our sample size, all data exclusions, and all measures in the study. The preregistration, analysis code, and study materials, can be found on the article's GitHub repository (https://github.com/stefanvermeent/liss_ef_2024). The newly collected data will be available for other researchers in the LISS data archive, after signing a data use agreement (https://lissdata.nl). For more information on the LISS variables used, see the article's GitHub repository (https://github.com/stefanvermeent/liss_ef_2024). Our study complies with Level 2 of the Transparency and Openness Promotion (TOP) guidelines.

## 3.3 Results

#### Structural equation model fit

As preregistered, we started with optimizing fit in structural equation models for each DDM parameter separately before combining them all into a single model. For drift rates, we selected the model containing a general processing speed factor (loading on all drift rates) and a common EF factor (loading only on drift rates of incongruent and switch conditions). An alternative model with a separate inhibition and attention shifting factor did not converge. The model with a common EF factor provided a good fit to the data (CFI = 0.98, RMSEA = 0.04 [0.03, 0.05]). However, the loadings of the common EF factor were small and in opposing directions, and the latent factor's residual variance was non-significant. Therefore, we did not estimate associations between adversity and common EF.

For both boundary separation and non-decision time, we selected a model containing only a general factor (loading on all boundary separations/non-decision times), which provided a good fit to the data (boundary separation: CFI = 0.98, RMSEA = 0.05 [0.04, 0.06]; non-decision time: CFI = 0.98, RMSEA = 0.04 [0.03, 0.05]). The two preregistered models specifying additional latent factors loading only on switching/incongruent conditions did not converge.

The final model combining all three DDM parameters provided a good fit, although it required the addition of covariances between residual variances of manifest DDM parameters of the same task (reflecting shared method variance), CFI = 0.96, RMSEA = 0.04 [0.04, 0.05]. See Figure 3.3 for a visualization and Table 3.3 for fit statistics of all models.

**Table 3.3.** Fit statistics of all preregistered structural equation models.

| Model | Chi square | Robust CFI | Robust RMSEA |
|---|---|---|---|
| **Measurement models** | | | |
| Drift rate model 1 | Model did not converge | | |
| Drift rate model 1 | 125.42 (45), p < .001 | 0.98 | 0.04 [0.03, 0.05] |
| Boundary separation model 1 | Model did not converge | | |
| Boundary separation model 2 | Model did not converge | | |
| Boundary separation model 1 | 165.37 (50), p < .001 | 0.98 | 0.05 [0.04, 0.06] |
| Non-decision time model 1 | Model did not converge | | |
| Non-decision time model 2 | Model did not converge | | |
| Non-decision time model 3 | 120.05 (50), p < .001 | 0.98 | 0.04 [0.03, 0.05] |
| Full measurement model | 1249.81 (445), p < .001 | 0.96 | 0.04 [0.04, 0.05] |
| **Structural models** | | | |
| Primary deprivation model | 1564.34 (599), p < .001 | 0.95 | 0.04 [0.04, 0.04] |
| Secondary deprivation model | 1625.91 (661), p < .001 | 0.96 | 0.04 [0.04, 0.04] |
| Primary threat model | 1309.71 (599), p < .001 | 0.95 | 0.04 [0.04, 0.04] |
| Secondary threat model | 1361.94 (602), p < .001 | 0.95 | 0.04 [0.04, 0.05] |

*Note:* for each DDM parameter, measurement model 1 contained a task-general latent factor as well as separate latent factors for inhibition and shifting. Measurement model 2 contained a task-general latent factor as well as a common EF factor. Measurement model 3 contained only a task-general latent factor.

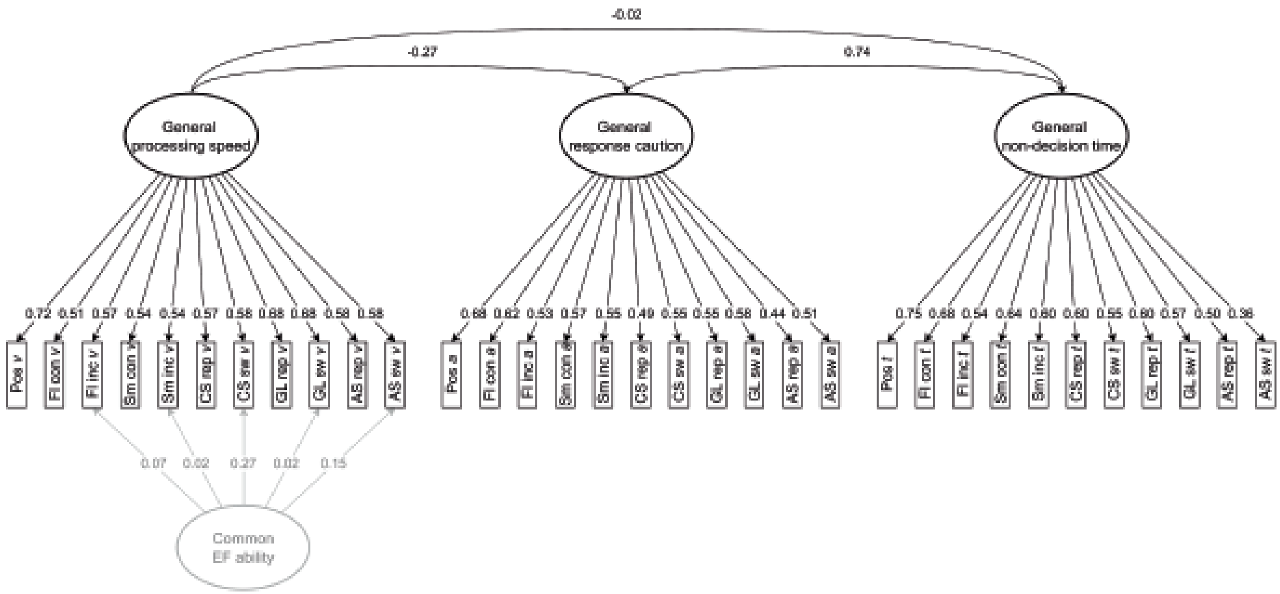



**Figure 3.3.** Final measurement model of the Drift Diffusion parameters. Rectangles depict manifest parameter estimates, elipses depict latent factors, unidirectional arrows depict factor loadings, and bidirectional arrows depict correlations. Latent factors with non-significant residual variances are greyed out. *a* = manifest boundary separation, AS = Attention-shifting task, CS = Color-shape task, GL = Global-local task, Fl = Flanker task, Sm = Simon task, *t* = manifest non-decision time, *v* = manifest drift rate.

## Preregistered analyses

Our primary analyses examined the association of exposure to material deprivation and threat in adulthood with task-general processing speed, task-general response caution, and task-general speed of stimulus encoding and/or response execution. As the structural equation model analyses reported above showed that EF-specific factors had non-significant variances, we retained them in the model but did not include them as dependent variables in subsequent analyses.

### Material deprivation in adulthood

Exposure to material deprivation in adulthood was not associated with either general processing speed ($\beta$ = −0.08, $p$ = .063), task-general response caution ($\beta$ = 0.06, $p$ = .155), or task-general non-decision time ($\beta$ = 0.05, $p$ = .153) (see Figure 3.4). In addition, none of the effects fell within the region of practical equivalence (all $p$s ≥ .092). The effects in the secondary model (including expossure to threat in adulthood as a confounder) were comparable.

### Threat in adulthood

Exposure to threat in adulthood was negatively associated with task-general processing speed ($\beta$ = −0.10, $p$ = .036) (see Figure 3.4). People who reported more exposure to threat in adulthood processed information more slowly across tasks. Exposure to threat in adulthood was not associated with task-general response caution ($\beta$ = 0.04, $p$ = .355), or task-general non-decision time ($\beta$ = −0.04, $p$ = .333). The effects in the secondary model (excluding material deprivation in adulthood as a confounder) were comparable, although the association with task-general non-decision time was practically equivalent in the secondary model.

## Non-preregistered analyses

### Associations with childhood adversity

In the first set of non-preregistered analyses, we analyzed how childhood exposure to threat and material deprivation are associated with task-general and ability-specific cognitive processes. The rational for these analyses is that we observed significant indirect associations between childhood adversity and general speed of processing in our preregistered analyses. Here, we analyzed *direct* effects of childhood threat and material deprivation (i.e., without including exposure to adversity in adulthood). We controlled for sex (but not age and education) and in the case of threat we also controlled for material deprivation, but not the other way around (similarly to the preregistered analyses).

Childhood exposure to material deprivation was negatively associated with general processing speed, $\beta$ = −0.24, SE = 0.04, 95% CI = [−0.31, −0.17], $p$ < .001 (Figure 3.5). People who reported higher levels of childhood exposure to material deprivation processed information more slowly across tasks. In addition, childhood exposure to material deprivation was positively associated with non-decision time, $\beta$ = 0.1, SE =

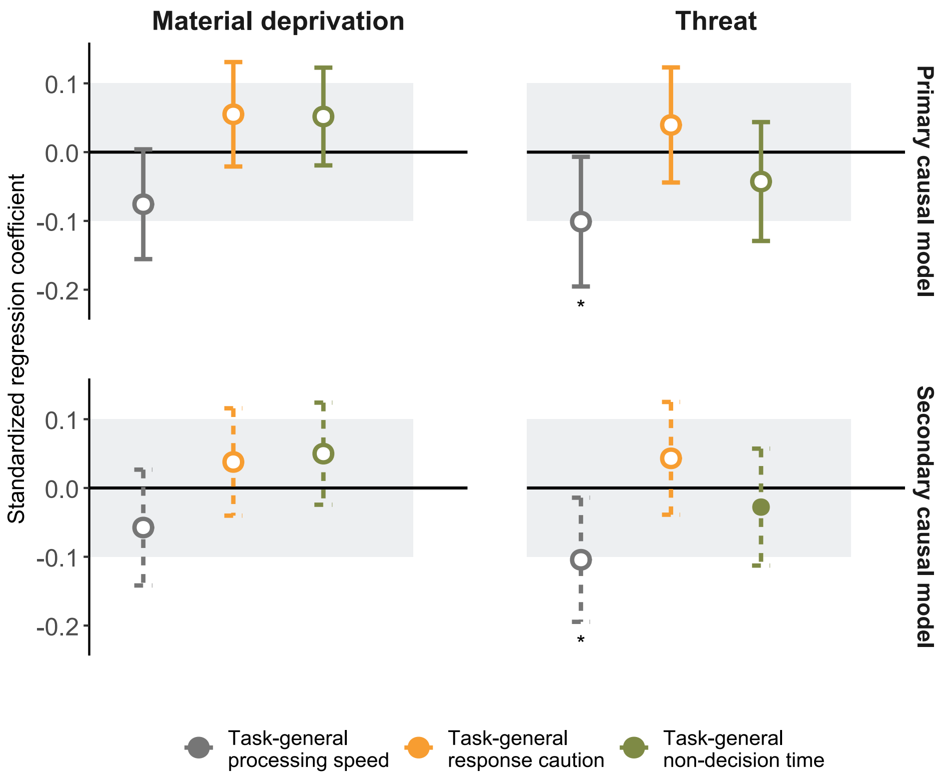



**Figure 3.4.** Standardized (preregistered) associations between exposure to adversity in adulthood with task-general processing speed, response caution, and non-decision time. The left panel depicts effects of material deprivation and the right panel depicts effects of threat. The top row depicts effects of the primary statistical model (following from our main Directed Acyclic Graph), and the bottom row depicts effects of the secondary statistical model (see Figure 3.2). For material deprivation in adulthood, the primary statistical model included age, education, and childhood adversity as confounders, whereas the secondary model additionally included threat in adulthood as a confounder. For threat in adulthood, the primary statistial model included age, childhood adversity, and material deprivation in adulthood, whereas the secondary statistical model did not include material deprivation in adulthood as a confounder. The gray area reflects the preregistered area of practical equivalence. Effects depicted with a solid point are practically equivalent, and effects depicted with a hollow point are not practically equivalent. Standard errors represent 95% confidence intervals. Statistical significance (tested against zero) is indicated with asterisks; * $p < .05$, ** $p < .01$, *** $p < .001$.

0.04, 95% CI = [0.02, 0.18], $p = .015$. People who reported higher levels of childhood exposure to material deprivation took more time to encode stimulus information and/or execute responses across tasks. Childhood exposure to material deprivation was not associated with general response caution. Equivalence tests revealed that none of these associations fell within the region of practical equivalence.

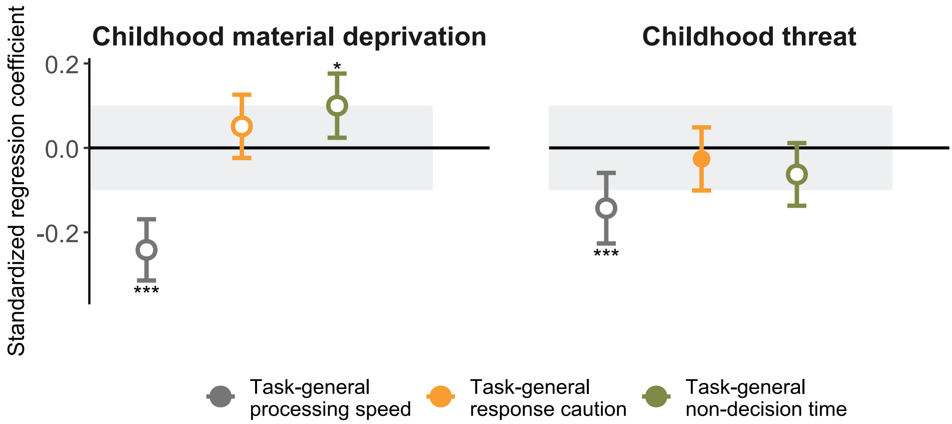



**Figure 3.5.** Standardized (non-preregistered) associations between childhood exposure to adversity with task-general processing speed, response caution, and non-decision time. The left panel depicts effects of material deprivation and the right panel depicts effects of threat. The gray area reflects the preregistered area of practical equivalence. Effects depicted with a solid point are practically equivalent, and effects depicted with a hollow point are not practically equivalent. Standard errors represent 95% confidence intervals. Statistical significance (tested against zero) is indicated with asterisks; * $p < .05$, ** $p < .01$, *** $p < .001$.

For childhood exposure to threat, we found a significant negative association with general processing speed, $\beta = -0.14$, SE = 0.04, 95% CI = [−0.23, −0.06], $p < .001$. People who reported higher levels of childhood exposure to threat processed information more slowly across tasks. We did not find significant associations between childhood exposure to threat and general response caution or general non-decision time. Of these associations, equivalence tests revealed that the association with general response caution fell within the region of practical equivalence ($p = .026$).

### Associations with task-specific drift rates

In the second set of non-preregistered analyses, we analyzed how adversity exposure during childhood and adulthood were associated with task-specific residual variances. The rationale for these analyses was that while we did not find coherent latent EF factors, all tasks showed significant residual variance after accounting for general speed of processing. This implies that, after accounting for general speed of processing, performance is further determined by unique features of individual tasks (i.e., method variance), which may be associated with adversity exposure. For this analysis, we were mostly interested in investigating the extent to which task-specific effects reflect EF abilities or task-specific processing, and therefore only focused on drift rates.

We first attempted to fit a model in which all task-specific drift rates were allowed to covary (i.e., across inhibition and attention-shifting tasks). However, this model did not converge. A model that only estimated covariances between task-specific drift rates of inhibition tasks and attention shifting tasks separately provided a good fit to the data (CFI = 0.99, RMSEA = 0.03 [0.02, 0.04]). This suggests that task-specific drift rates correlated among inhibition tasks and among attention-shifting tasks, but not between inhibition and attention-shifting tasks.

Figure 3.6 shows latent correlations between task-specific drift rates after accounting for task-general processing speed. The drift rates of repeat and switch conditions within each attention shifting task correlated moderately to strongly, while all correlations between tasks were low. This suggests that task-specific drift rates of attention shifting tasks captured features of the tasks that were not shared between tasks. In contrast, all correlations between task-specific drift rates of inhibition tasks were low, even among drift rates of the same task. Thus, task-specific drift rate did provide consistent measures of specific EF abilities that are shared across tasks.

Next, we explored the associations of task-specific drift rates with exposure to material deprivation and threat in both childhood and adulthood, after accounting for task-general processing speed (Figure 3.7). Exposure to material deprivation in adulthood was not associated with task-specific drift rates of any of the tasks. Exposure to threat in adulthood was negatively associated all task-specific drift rates in both the repeat and switch condition of the Color-shape task (repeat: $\beta = -0.14$, SE = 0.05, 95%

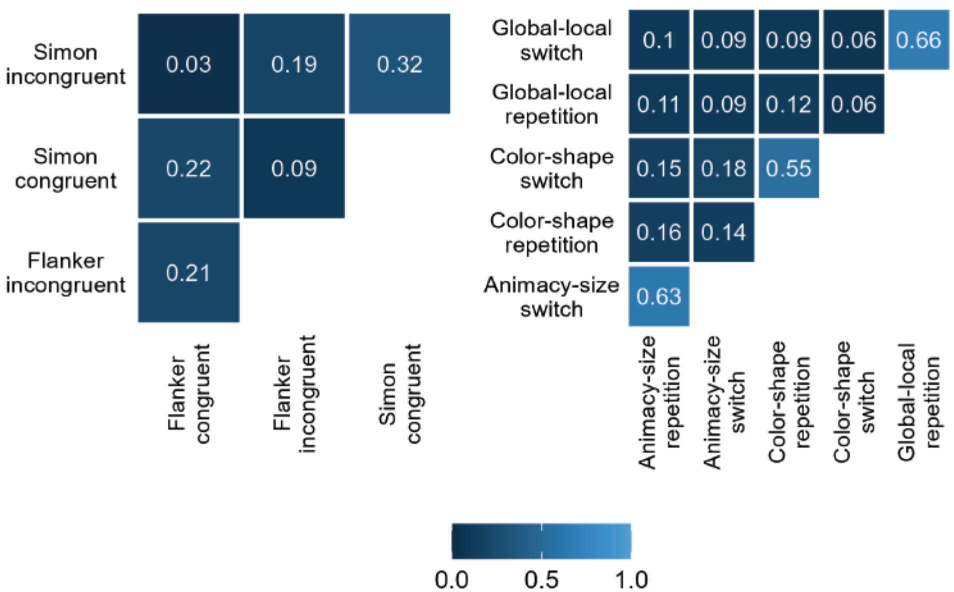



**Figure 3.6.** Correlations between task-specific drift rates after accounting for task-general processing speed.

CI = [−0.23, −0.05], $p$ = .025; switch: $\beta$ = −0.13, SE = 0.05, 95% CI = [−0.22, −0.04], $p$ = .025). None of the other associations were significant.

Childhood exposure to material deprivation was negatively associated with all task-specific drift rates, with effect sizes ranging from $\beta$ = −0.08 for the congruent condition of the Simon Task to $\beta$ = −0.27 for the incongruent condition of the Global-local Task. Childhood exposure to threat was negatively associated with all but the task-specific drift rates of the congruent condition of the Flanker and the Simon task. The effect sizes of significant effects ranged from $\beta$ = −0.06 for the incongruent condition of the Simon Task to $\beta$ = −0.16 for the switch condition of the Posner Task.

To summarize, we did not find evidence that exposure to threat or material deprivation (either in childhood or adulthood) was associated to specific inhibition or attention-shifting ability. Rather, we mostly found associations with task-general processes. We did find associations between childhood adversity and task-specific drift rates. However, correlations among task-specific drift rates were low, even for tasks thought to measure the same EF ability, suggesting that they reflect method variance rather than specific EF abilities.

## 3.4 Discussion

We investigated associations between exposure to two types of adversity in adulthood —material deprivation and threat—with inhibition and attention shifting ability. Participants completed two inhibition tasks, three attention shifting tasks, and one basic processing speed task. First, we used DDM to separate raw performance into three distinct cognitive processes: (1) speed of evidence accumulation (drift rate), (2) response caution (boundary separation), and (3) speed of stimulus encoding and response execution (non-decision time). Finally, we used structural equation modeling to separate variance in each cognitive process into a task-general factor (shared across all tasks) and ability-specific factors.

### Main findings

People with more exposure to threat in adulthood—but not material deprivation—showed lower general processing speed. This is consistent with predictions from deficit frameworks (Sheridan & McLaughlin, 2014; Tucker-Drob, 2013). After accounting for general processing speed, there was no remaining variance that could be attributed to either inhibition ability or attention shifting ability. Additionally, the associations between material deprivation and threat in adulthood with both response caution and non-decision time were neither significantly different from zero, nor practically equivalent to zero. In other words, from these data, we were neither able to conclude that people with more exposure to adversity differed in their level of response caution and non-decision time, nor that these processes were similar to people with lower levels of adversity exposure.

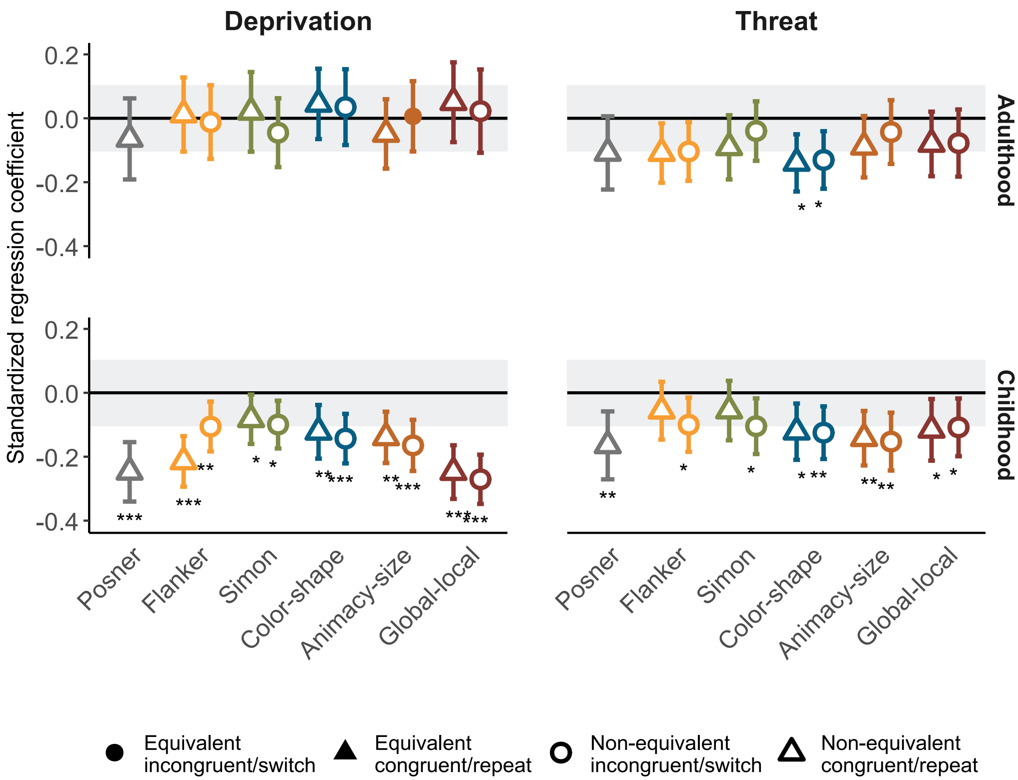



**Figure 3.7.** Standardized (non-preregistered) associations between adversity and task-specific drift rates. The effects are split out over adversity type (columns) and timing of adversity (rows). The gray area reflects the area of practical equivalence. Triangles depict effects for repeat/congruent conditions, and circles reflect effects for switch/incongruent conditions. Effects depicted with a solid point are practically equivalent, and effects depicted with a hollow point are not practically equivalent. Standard errors represent 95% confidence intervals. Statistical significance (tested against zero) is indicated with asterisks; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We found that variance in EF task performance was mostly explained by a task-general factor, not by ability-specific factors. This finding aligns with previous research applying DDM to EF tasks. Three recent studies converged on the similar conclusion that performance across different EF tasks is best explained by a task-general drift rate factor (Frischkorn et al., 2019; Hedge et al., 2022; Löffler et al., 2024; Weigard et al., 2021). One study found that the task-general drift rate factor fully reflected basic processing speed, rather than common EF processes (Löffler et al., 2024). Similarly, another study including several inhibition tasks found that shared variance was mostly explained by basic processing speed and strategy differences, not by inhibition ability (Hedge et al., 2022). It is therefore not surprising that we were unable to estimate ability-specific EF factors after accounting for general processing. These issues tie into several other criticisms of common EF tasks in the cognitive psychology literature—such as low test-retest reliability—which should be considered when using these tasks to test associations with adversity (Bastian et al., 2020).

Our results also align with research showing that associations between adversity and raw performance (e.g., response times, accuracy) on EF tasks are mostly task-general. For instance, across three large cohorts, youth from more disadvantaged backgrounds had lower task-general raw performance across a wide range of cognitive tasks. After controlling for task-general effects, performance on specific intelligence tasks was largely practically equivalent to zero, and performance on some specific EF tasks even appeared enhanced (Bignardi et al., 2024). Finally, previous research applying DDM in the context of adversity found similar (but not identical) results as reported here. Youth with more exposure to threat (but not material deprivation) had slower general processing speed across three EF tasks and a basic processing speed task, while most specific abilities remained intact (Vermeent et al., 2024). In contrast to the findings reported here, these youth also showed more response caution. It is important to highlight that both studies measured specific EF abilities with only a single task. This makes it unclear to which extent the task-specific effects they observed reflected EF abilities (see also our non-preregistered findings discussed below).

**Non-preregistered findings**

Additional non-preregistered analyses showed that childhood exposure to both material deprivation and threat were associated with slower general processing speed. The effect sizes of these associations were larger than those of recent adversity exposure. This finding is striking considering the average age in our sample was 39 years. This is consistent with previous literature suggesting that adverse experiences early in life have a lasting effect on brain development (Nelson et al., 2020; Nelson & Gabard-Durnam, 2020; Shonkoff et al., 2012). For instance, both experiences of malnutrition and exposure to stress (e.g., as a result of threat) have been linked to structural changes in brain networks involved in cognitive functioning (Algarin et al., 2017; Polavarapu & Hasbani, 2017; Rebello et al., 2018).

A second important non-preregistered finding is that after controlling for task-general processing, adversity exposure was negatively associated with several task-specific drift rates, mostly relating to childhood exposure to threat and deprivation. Importantly, the correlations among task-specific drift rates were all small—with the exception of correlations among switch and repeat conditions of the same attention shifting tasks. This suggests that these negative associations should not be interpreted as lower ability-specific processing, relating to inhibition and attention shifting ability. Instead, it appears to relate to information processing that is much more specific to individual tasks. This is consistent with literature showing that a substantial part of the variance in performance on cognitive task consists of method variance resulting from the nature of the task itself, rather than what the task is designed to measure (Barkley, 2012; Rey-Mermet, 2024; Schubert et al., 2016).

**Implications for adversity research and interventions**

Our results have two implications for adversity research and interventions, both of which warrant caution against interpreting raw performance on individual EF tasks. First, adversity appears to be associated mostly with task-general processes, rather than specific abilities. This contrasts with common inferences in both the deficit and adaptation literature, which tend to explain performance differences in terms of specific EF abilities, such as inhibition (Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015; Noble et al., 2005) and attention shifting (Fields et al., 2021; Howard et al., 2020; Mittal et al., 2015; Nweze et al., 2021; Young et al., 2022). Moreover, our results show that accounting for task-general processing does not guarantees that task-specific associations with adversity reflect differences in EF. Instead, negative associations with adversity appear to be partly driven by specific features of individual tasks. A general recommendation for adversity research is therefore that studies should include (1) multiple EF tasks to quantify task-general processing speed, and (2) ideally two or more tasks measuring the same EF ability to quantify ability-specific processing. DDM and SEM can then be used to account for task-general and task-specific processes. For interventions, our results imply that individual EF tasks, or batteries of EF tasks each of which measures a different ability, have limited value as screening tools; and, that training performance on such tasks is unlikely to transfer to broader, sustainable changes in EF abilities.

Our findings complicate research that aims to understand how EF abilities develop in adverse environments. We agree with others that results like ours do not necessarily mean that EF abilities do not exist (e.g., Löffler et al., 2024), or that adversity is not associated with differences in EF. Rather, it points to issues with how we *measure* EF abilities. It is difficult to provide a straightforward solution for this problem, which is still a much-debated topic in cognitive psychology (Bastian et al., 2020). Part of the solution may lie in adjusting the content of EF tasks, for instance, by using real-world content rather than abstract content (Young et al., 2022). By increasing familiarity, relevance, and/or valence, such content might increase the engagement of EF abilities (Niebaum & Munakata, 2023). Another solution could be to use adaptive testing procedures, which appear to improve estimation of EF abilities (Draheim et al., 2021). However, more fundamental changes might be necessary as well, such as developing new paradigms that align more closely to people's everyday experiences, goals, and challenges (Doebel, 2020; Miller-Cotto et al., 2022; Niebaum & Munakata, 2023).

**Strengths, limitations, and future directions**

Our study has three main strengths. First, we measured inhibition and attention shifting ability using multiple tasks for each EF ability. This approach provided two advantages: (1) we could distinguish between general versus specific processes, and (2) we could measure EF abilities without task-specific measurement error. Second, we used

DDM to distinguish between evidence accumulation (EF ability, general speed of processing), response caution, and non-decision time (stimulus encoding, response execution). The DDM allowed us to better understand the cognitive processes driving performance in people with more exposure to adversity. Third, we included a large, representative sample of the Dutch population sample with sufficient variance on key adversity measures.

Our study has three main limitations. First, although we included confounding variables based on explicit, theory-guided causal assumptions, the cross-sectional and partly retrospective nature of our data prohibit us from establishing true causal relationships. Second, due to time constraints, we only included two inhibition tasks and three attention shifting tasks, limiting our ability to model shared variance between them. Future studies could include more EF tasks, and ideally include several basic processing speed tasks. Third, participants completed the cognitive tasks online in their own home. Despite our controlling for environmental factors that may have disrupted task performance, the home setting may have reduced the reliability of performance measures.

We envision three future directions. First, because adversity tends to be associated with task-general processes, it will be important to better understand the nature of these processes. To the extent that task-general drift rates reflect general processing speed, it may partly result from lower white matter tract integrity (Fuhrmann et al., 2020; Kievit et al., 2016). Such associations with structural brain differences may account for the finding that task-general drift rate is relatively stable over time (Schubert et al., 2016; Weigard & Sripada, 2021). Yet, task-general drift rate may be associated with other factors such as effort, fatigue, or hunger (Weigard & Sripada, 2021), which may play a bigger role for people from adverse environments (Brose et al., 2012; Schwabe et al., 2013; Sliwinski et al., 2006). Thus, future studies may include (a combination of) brain measures or measures of mental states to better understand inter- and intrapersonal differences in task-general drift rates.

Second, future research should investigate why adversity is negatively associated with task-specific drift rates. Our results suggest that task-specific drift rates mostly capture method variance rather than specific EF abilities. We currently do not have a clear explanation for their negative associations with adversity. One possible explanation might be that the associations reflect people's difficulty with processing abstract content. Previous research found that youth with more exposure to adversity performed lower on a working memory updating task (compared with youth with fewer adversity exposures) when task stimuli were abstract geometric shapes. However, their performance equalized when the task included real-world content (Young et al., 2022). All tasks in our study included abstract content, and each study used different stimuli, which may account for the negative associations with adversity and the lack of correlations between task-specific drift rates. However, if so, this would not explain the low correlations between congruent and incongruent conditions of both the Flanker and Simon task, as these conditions do include the same stimuli. Nevertheless, research could more systematically test associations between adversity and processing as a function of task content.

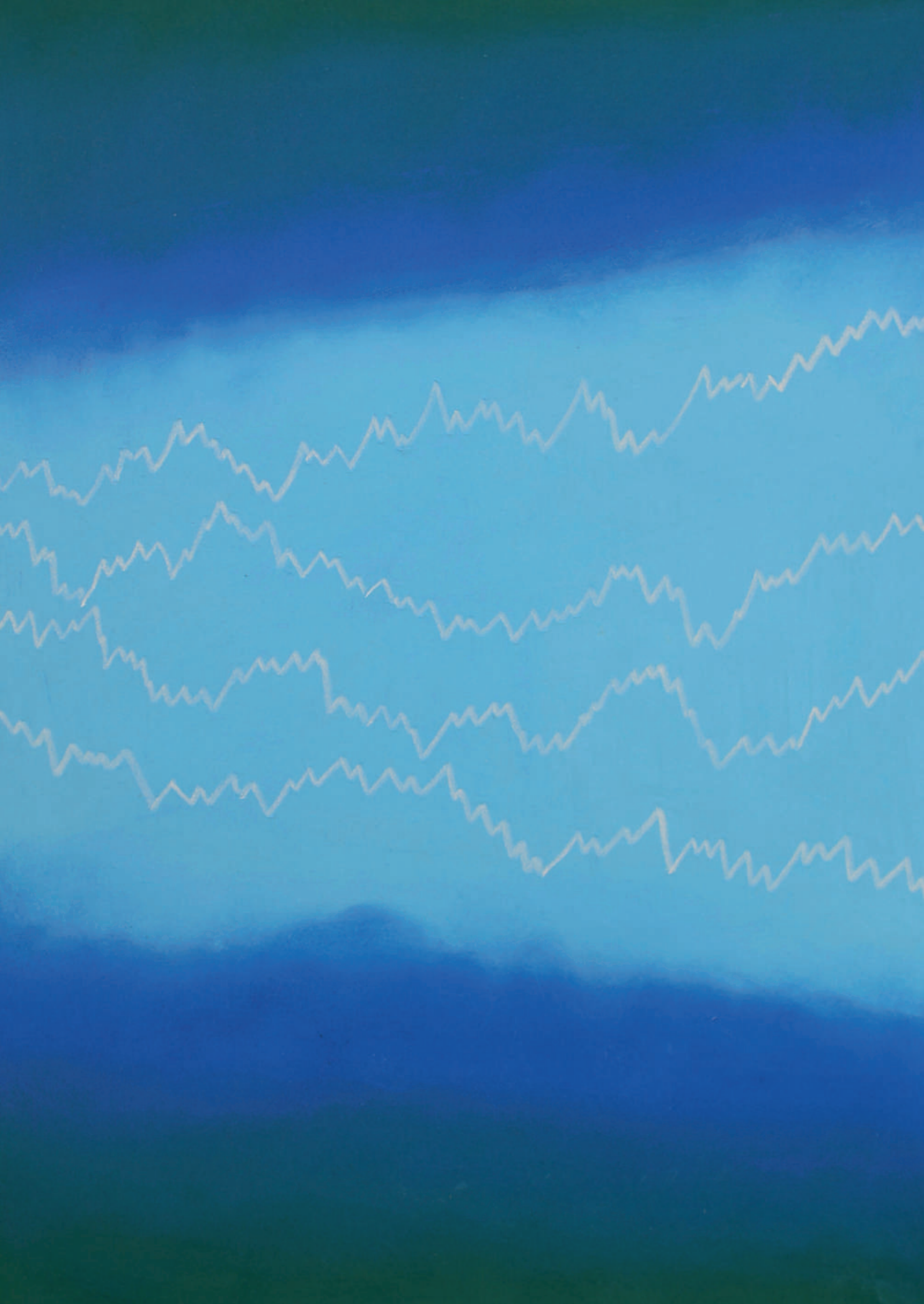
Third, future research could design preregistered, well-powered analyses to tease apart the unique effects of childhood adversity and recent adversity on cognitive performance, as well as their combined effects. Our understanding of how cognitive outcomes later in life depend on the developmental timing of adversity exposure is still limited (Frankenhuis, Young, et al., 2020). In this study, we assumed that recent adversity exposure mediates the effect of childhood adversity exposure on cognitive processes, as people who experienced adversity early in life are more likely to also experience adversity later in life (e.g., due to systemic constraints; Hazel et al. (2008)). Our study focused on adversity in adulthood, the average or summed adversity exposure over the past few years. Future research should distinguish such recent adversity from acute stress. For instance, some research suggests that childhood adversity enhances specific EF abilities, but only under situations of acute stress (Mittal et al., 2015; Young et al., 2018). Thus, future research should tease apart effects of adversity during childhood and adulthood and study how their effects are moderated by acute stress.

### Constraints on generality

Our research question concerned the association between adversity exposure and performance on EF tasks, using cognitive modeling to investigate whether these associations are driven by task-general processing speed or specific EF abilities. Our target population was Dutch adults with a broad range of adversity exposures. The LISS panel ensures representation of the Dutch adult population in terms of—among other factors—age, education, and socioeconomic status. Our findings may generalize to adults in other Western populations, but not necessarily to adults in non-Western populations (Nketia et al., 2024). We used standardized inhibition and attention-shifting tasks that are often used in cognitive psychology. The findings should generalize to other tasks used to measure these abilities, as long as they rely on speeded responses. All tasks included abstract stimuli (e.g., geometric shapes, words). Results are likely to be different when using more real-world stimuli, or when changing the tasks in other ways to align them with people's lived experiences (Doebel, 2020; Miller-Cotto et al., 2022; Young et al., 2022). Participants completed the study online in their home environment. To reproduce the effects, it is important to carefully instruct participants to limit environmental distractions and complete the study in full-screen mode. The results should generalize to more standardized settings (e.g., the lab), but may not generalize to public settings outside of the home environment.

### Conclusion

Adversity research has made important steps in recent years, revealing how exposure to adversity may impair certain EF abilities while enhancing others. Further progress in this field hinges on our capacity to accurately measure specific EF abilities. We show, consistent with previous research, that this is a more challenging task than typically assumed. As a result, adversity researchers likely overestimate the strength of associations between adversity and EF abilities when only analyzing raw performance. We present a way forward by building on a combination of cognitive modeling and structural equation modeling. Embracing these techniques, and revising theories in light of their findings, will enhance our understanding of how exposure to adversity shapes EF.



# Chapter 4

Childhood adversity is not associated with lowered inhibition, but lower perceptual processing: A Drift Diffusion Model analysis

**This chapter is based on**
Vermeent, S., Young, E.S., van Gelder, J.-L., & Frankenhuis, W.E. (2024). Childhood adversity is not associated with lowered inhibition, but slower perceptual processing: A Drift Diffusion Model analysis. *Cognitive Development, 71*, 101479. https://doi.org/10.1016/j.cogdev.2024.101479

## 4.0 Abstract

It is well-established that individuals who grew up in adverse conditions tend to be slower on the Flanker Task. This finding is typically interpreted to reflect difficulty inhibiting distractions. However, it might result from slower general cognitive processes (e.g., reduced general processing speed), rather than the specific ability of inhibition. We used Drift Diffusion Modeling in three online studies (total N = 1560) with young adults to understand associations of adversity with Flanker performance. We find no associations between exposure to violence and unpredictability with inhibition. Yet, although mixed, violence and unpredictability exposure were associated with lower strength of perceptual input—how well someone can process target and distractor information alike. Finally, people with lower strength of perceptual input processed information more holistically, focusing less on details. Thus, lowered Flanker performance does not necessarily imply lowered inhibition ability. Cognitive modeling reveals a different picture of abilities in adverse conditions as opposed to analyses based on raw performance.

**Author contributions**

All authors were involved in conceptualizing the study. SV coordinated the data collection and analyzed the data, and wrote the first draft of the manuscript. All authors provided feedback on the manuscript.

## 4.1 Introduction

The predominant view in developmental psychology is that exposure to adversity—defined as prolonged exposure to intense stress—impairs cognitive abilities. This view is supported by decades of research showing that people living in high-adversity contexts tend to score lower on a variety of cognitive tests (Hackman et al., 2010; Ursache & Noble, 2016a). Recent adaptation-based perspectives, however, have argued that people from adversity may also develop intact, or enhanced, abilities for solving problems in high-adversity contexts (Ellis et al., 2017; Frankenhuis & Weerth, 2013). Adaptation- and deficit-based perspectives are considered complementary. For instance, adversity may impair some cognitive processes, yet enhance others. Despite their compatibility, few studies have investigated how the interplay of impaired and enhanced abilities shapes performance. Across three preregistered online experiments, we used cognitive modeling to derive a process-level understanding of the association between childhood adversity and performance in the Flanker task, a popular measure of cognitive control (Ridderinkhof et al., 2021).

**Attention in adverse conditions**

It is well-established that early-life adversity is associated with deficits in the ability to inhibit distracting, goal-irrelevant information (Hackman et al., 2010; Ursache & Noble, 2016a). One of the leading paradigms in this literature is the Flanker task (B. A. Eriksen & Eriksen, 1974). On this task, participants typically see five arrows in a horizontal orientation, and are asked to indicate the direction of the central arrow. The flanking arrows point in the opposite direction on half of the trials, leading to interference that participants must inhibit. Slower performance in the Flanker task has been documented for children and adults with more environmental unpredictability (Fields et al., 2021; Mittal et al., 2015). These findings are typically interpreted as indicating a deficit in the ability to inhibit distractions.

Similar associations have been documented for factors that increase the risk of adversity exposure, such as lower socioeconomic status (SES; Farah et al., 2006; Mezzacappa, 2004; Noble et al., 2005). Although people living in low-SES conditions experience more adversity, on average, we do not regard low SES itself as a form of adversity. First, SES and adversity can affect cognitive abilities through different mechanisms (e.g., education versus physiological stress). Second, people with low SES have diverse experiences, both positive and negative, even if adversity is more common in this group.

Some recent studies suggest that growing up in adversity may also be associated with improved abilities such as attention shifting (Fields et al., 2021; Mittal et al., 2015; Young et al., 2022; but see Mezzacappa, 2004; Nweze et al., 2021). Some studies found deficit patterns on inhibition tasks alongside enhancements on other aspects of attention within the same participants. For example, one study found that young adults with more childhood unpredictability committed more errors on an Antisaccade task (a measure of inhibition), but more efficiently switched their attention between tasks on an attention shifting task (Mittal et al., 2015). Similarly, children with more caregiver switches (an indicator of unpredictability) experienced more interference in the Flanker task (based on RTs), but outperformed children with fewer caregiver switches on shifting their attention between different task goals (Fields et al., 2021).

Performance on attention tasks could reflect developmental adaptation to adverse environments (Blair & Raver, 2012; D'angiulli, Lipina, et al., 2012; Frankenhuis, Young, et al., 2020; Mittal et al., 2015). In unpredictable or threatening conditions, the ability to detect salient peripheral information (e.g., distant noises or approaching individuals) could help to more quickly detect and act on potential threats. Over time, cognitive adaptations to such conditions could result in a general tendency to use a more diffuse scope of attention, leading to an enhanced ability to keep track of the broader environment. In line with this hypothesis, people with lower SES respond more strongly to auditory distractors (Giuliano et al., 2018; Hao & Hu, 2024; Stevens et al., 2009) and are faster to orient their attention to peripheral visual information (Mezzacappa, 2004). While potentially adaptive, a more diffuse scope of attention could come at the cost of lowered ability to ignore irrelevant distractors. This could compromise longer-term goal-directed behavior, especially in chaotic environments (e.g., a noisy classroom or a busy street).

Thus, lowered performance on tasks like the Flanker task could reflect either a cognitive impairment or a difference in attentional strategies. Distinguishing between these two possibilities is challenging for two reasons. First, few studies in the adversity literature have measured performance differences on different attention tasks within the same individual (Mezzacappa, 2004; Mittal et al., 2015). Thus, it is unclear whether lowered inhibition is related to enhanced processing of peripheral information in people from adverse backgrounds. Second, performance on inhibition tasks is—beyond the ability to inhibit distractors—also influenced by other factors, such as a person's general processing speed and response caution (Hedge et al., 2022; Löffler et al., 2024). This means that lowered performance on inhibition tasks does not necessarily reflect inhibition difficulties. In other words, we should consider cognitive processes other than ability when drawing inferences based on inhibition tasks.

### Using Drift Diffusion Modeling to estimate attention and processing styles

An important issue, therefore, is that several processes are involved in performance in the Flanker task, and standard assessments using raw performance measures (response times, accuracy rates) mostly fail to distinguish between them. For example, performance differences in the Flanker task could indicate that someone experiences more (or less) distractor interference, generally processes less (or more) efficiently, or responds with less (or more) caution. To understand how adversity affects performance, we need to be able to separate the difference processes that make up performance.

Formal cognitive models such as the Drift Diffusion Model (DDM; Forstmann et al., 2016; Ratcliff et al., 2015; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Wagenmakers, 2009) provide a potential solution. The DDM estimates explicitly models the cognitive processes underlying the decision-making (See Figure 4.1a). It represents decision-making on binary decision-making tasks as a process in which people accumulate information until one response is sufficiently favored over the other. These two response options are represented as opposing boundaries. One boundary corresponds to the correct response and the other to the incorrect response (note that in some research designs, the boundaries may be coded as the two choice options instead, for example, when the question is whether people classify a certain class of stimuli (e.g., angry faces) more efficiently than another class of stimuli (e.g., happy faces)). When the accumulated information reaches one of the two boundaries, the corresponding response is executed.

The DDM translates trial-level response times (RTs) and accuracy into three distinct cognitive processes. The speed of information accumulation is captured in a parameter called the *drift rate*. Higher drift rates are associated with faster responses and higher mean accuracy. Response caution is modeled through the *boundary separation*; that is, the width between the two boundaries. Larger boundary separation is associated with larger RTs and higher accuracy (i.e., sacrificing speed to increase accuracy). *Non-decision time* represents the time it takes to prepare for the task at the start of the trial (before information accumulation starts) and the time it takes to execute a response (after a response boundary has been reached). Longer non-decision times are associated with larger RTs, without influencing accuracy. Finally, the *starting point* represents a potential bias towards one of two responses, with a biased decision-making process starting closer to one boundary relative to the other boundary.

The Shrinking Spotlight (SSP) model is an extension of the standard DDM to account for attention processes in the Flanker task (Grange, 2016; White et al., 2011, 2018; White & Curl, 2018). The SSP model assumes that attention resembles a spotlight that is normally distributed over the Flanker task arrows (with a particular starting *attentional width*). Over time, people narrow their attention down to the central arrow (at a rate defined by the *shrinking rate*), thereby gradually decreasing interference from irrelevant information (cf. C. W. Eriksen & St. James, 1986; see Figure 4.1b). Prior work has defined the amount of distractor *interference* by dividing the attentional width by the shrinking rate (White et al., 2018). People may experience less interference either by starting with a narrower attentional width, and/or by more rapidly shrinking their attention down to the target arrow. Finally, performance is also influenced by the *perceptual input* strength; that is, how well someone can process the arrows in general. Note that typical interpretations of lowered raw Flanker task performance are in terms of the amount of interference that someone experiences, and not in terms of the strength of perceptual input.

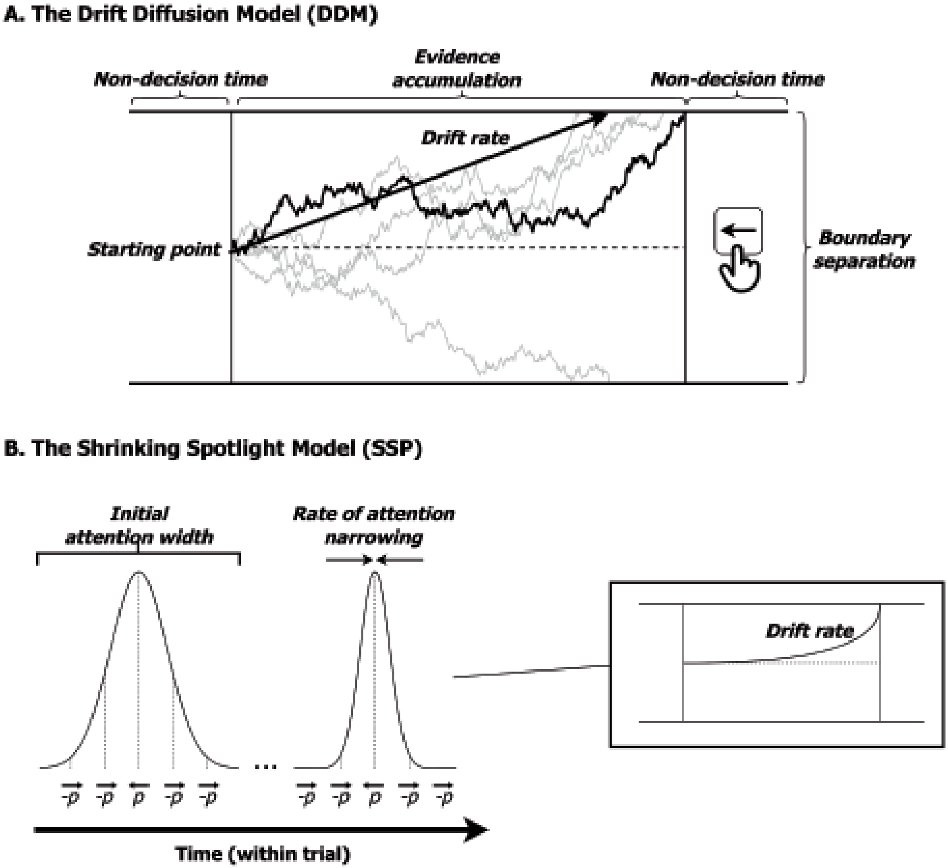



**Figure 4.1.** A visual overview of the Drift Diffusion Model (DDM) and the Shrinking Spotlight Model (SSP). Panel A: The DDM assumes that people go through three distinct stages on cognitive tasks with two response options. In a first preparation phase, they encode the stimuli. In a second decision phase, people accumulate evidence for both decisions (e.g., pressing left vs. right) until the decision boundary for either the correct or incorrect response is reached. Each jagged line represents this information accumulation process on a single trial. In a third execution phase, people execute the motor response. The model estimates four parameters reflecting distinct cognitive processes (printed in italic): (1) The drift rate represents the average rate of evidence accumulation and measures processing speed; (2) The non-decision time represents stimulus encoding and response exeuction; (3) The boundary separation represents the space between decision boundaries, and measures a person's level of response caution; (4) The starting point (not considered here) represents a potential bias towards a response, with a biased decision-making process starting relatively closer to one boundary. Panel B: The SSP is an extension of the standard DDM including additional parameters to capture attentional processes on the Flanker task. Each arrow provides a certain strength of perceptual input (*p*). On incongruent trials, the perceptual input of flanking arrows is coded negatively (-*p*). Attention is assumed to be normally distributed over the arrows with a certain attentional width. Over time, attention narrows down toward the central arrow at a rate determined by the shrinking rate, which gradually lowers interference from flanking arrows. The drift rate in the SSP model is the sum of each arrow's perceptual input multiplied by the attention allotted to each arrow. As attention for the flanking arrows decreases over time, the drift rate is assumed to increase over time (contrary to the standard DDM, which assumes a linear drift rate).

The DDM and SSP model share many assumptions. Both provide identical estimates of boundary separation and non-decision time. The main difference is the decision-making phase. The DDM assumes that the quality of the information is the same across the entire trial. In contrast, the SSP model assumes that the quality of the information improves across the trial, as attention becomes gradually focused more on the central arrow. The simpler assumption of the DDM makes the model broadly applicable, but less precise for conflict tasks. The SSP model is more precise, but applicable only to the Flanker task. Previous studies have successfully applied the DDM to Flanker task data (e.g., Löffler et al., 2024; Vermeent et al., 2024). However, unlike the DDM, the SSP model affords testing hypotheses about the association between adversity and attentional interference in the Flanker task. Specifically, we are interested in how childhood adversity is associated with both interference and the strength of perceptual input. Finally, the SSP model is one of several completing diffusion models developed to explain performance on conflict tasks (Hübner et al., 2010; Ulrich et al., 2015). We focus exclusively on the SSP model because it performs well with relatively few trials per participant relative to other conflict diffusion models (White et al., 2018).

**Overview of studies**

The overarching goal of our studies is to understand the attentional and processing styles that people develop in conditions of adversity. We focus on measures of exposure to violence and environmental unpredictability. Previous research shows that these two types of adversity are on the one hand associated with improved attention shifting and working memory updating (Fields et al., 2021; Mittal et al., 2015; Young et al., 2018, 2022), and on the other hand with lowered inhibition and working memory capacity (Fields et al., 2021; Mittal et al., 2015; Young et al., 2018). We conducted three online studies: one pilot study and two follow-up studies. Using cognitive modeling, we unpack Flanker task performance in comparison to other tasks that require externally focused attention (Pilot study), across visual processing manipulations (Study 1), and in terms of tendencies for holistic versus local processing (Study 2).

We used an incremental preregistration approach across studies (for all preregistrations, data, code and materials, see https://stefanvermeent.github.io/attention_project/). For each study, we preregistered confirmatory (i.e., hypothesis-driven) and exploratory analyses. The main text addresses the confirmatory analyses involving violence exposure and the exploratory analyses involving environmental unpredictability. We describe the other exploratory analyses in the supplemental materials (section 2). For an overview of all deviations from the preregistrations, see section 4 of the supplemental materials.

## 4.2 Pilot study

In the Pilot study, our goal was to understand how childhood adversity relates to performance on tasks with different attentional demands. Participants completed self-report measures of childhood adversity and three cognitive tasks (Flanker task, Cued Attention task, and Change Detection task). These tasks measured inhibition, attention for peripheral cues, and attention for subtle changes. In line with the idea that exposure to adversity may lead to a more diffuse scope of attention, we expected people with more violence exposure to be better at detecting peripheral stimuli and subtle changes. We expected this would result in a higher drift rate (faster speed of information accumulation) or shorter non-decision times (faster attention orientation, among other things), but not necessarily with differences in boundary separation (response caution). In contrast, we expected that participants with more violence exposure would be worse at ignoring distracting peripheral stimuli. We expected this would result in more experienced interference (as derived from the SSP model).

## Methods

### Participants

Participants were 565 people from the United States aged between 18 and 30 recruited on Prolific Academic (https://www.prolific.co) (See Table 4.1 for demographic data). The sample was balanced on sex. We used the MacArthur's ladder, included in Prolific's prescreening battery, for assessing perceived SES to ensure about half of the sample came from lower-SES backgrounds (which we defined as a score of 4 or below). Participants were eligible if they spoke fluent English and did not report color-blindness. We obtained ethical approval from the Ethics Review Board of the Faculty of Social & Behavioral Sciences of Utrecht University (FETC20-490).

**Table 4.1.** Demographic information for all studies.

| | Pilot study | Study 1 | Study 2 |
|---|---|---|---|
| **N** | 512 | 497 | 551 |
| **Mean age (SD)** | 24 (4) | 25 (3) | 26 (3) |
| **Sex (%)** | | | |
| Male | 49.4 | 49.3 | 49.9 |
| Female | 50.0 | 49.7 | 49.2 |
| Prefer not to say | 0.6 | 0.8 | 0.9 |
| Intersex | 0 | 0.2 | 0 |
| **Highest education (%)** | | | |
| Some high school | 1.2 | 1.2 | 1.6 |
| GED | 1.8 | 1.6 | 2.4 |
| High school diploma | 17.8 | 14.7 | 15.1 |
| Some college but no college degree | 32.6 | 27.6 | 25.2 |
| Associate's degree | 6.8 | 8.2 | 10.2 |
| Bachelor's or RN degree | 31.6 | 37.2 | 35.8 |
| Master's degree | 7.2 | 7.8 | 7.6 |
| Doctoral or law degree | 1.0 | 1.6 | 1.6 |
| Prefer not to say | 0 | 0 | 0.5 |
| **Social class (%)** | | | |
| Poor | 6.0 | 7.8 | 7.8 |
| Working class | 30.5 | 34.8 | 36.7 |
| Middle class | 46.1 | 40.2 | 39.2 |
| Upper-middle class | 16.4 | 16.1 | 14.5 |
| Upper class | 1.0 | 0.8 | 0.9 |
| Don't know/prefer not to say | 0 | 0.2 | 0.9 |

We conducted a power simulation using the *faux* package in R (DeBruine, 2021) to determine the minimally required number of participants for standardized regression coefficients of 0.10 and 0.15 (for details and simulation code, see https://stefanvermeent.github.io/attention_project/preregistrations/README.html). Power was > .80 for adversity x task condition interactions with $N$ = 450 or more. For a linear main effect, detecting an effect of $\beta$ = 0.15 with .90 power would require $N$ = 462. We sampled 550 participants, anticipating a final sample of ~500 after exclusions.

Prior to analyzing the data, we applied our preregistered exclusion criteria. First, we excluded participants who did not complete the full study; second, those who had incomplete data on any of the attention tasks; third, those who missed both attention check items; fourth, those who had suspicious response patterns (e.g., consistently endorsing high response options even though some items were reverse coded). Fifth, on a trial-level, we excluded any trials with reaction times < 250 ms or > 3500 ms (Ratcliff & Childers, 2015). Participants with more than 10 trials removed were completely excluded from the analyses. The final sample consisted of 512 participants.

**Procedure**

Participants completed the experiment on their own laptop or desktop computer. Participants could refrain from answering any of the questionnaire items and were prompted with a warning once per page in case of missing items.

After providing consent, participants completed three attention tasks. They were asked to move to a quiet room in the house, where they would be unlikely to be distracted by other people or outside noises. The order of the tasks was counterbalanced between subjects. At the onset of the first task, the experiment went into full-screen mode to limit distractions from other programs or browser tabs. The size of the task stimuli was controlled between subjects using the resize plugin in JsPsych (De Leeuw, 2015). Participants were asked to hold a credit card (or similarly sized card) up against the screen and to increase the size of a blue rectangle on the screen until it matched the size of the credit card. The stimulus display for each task was resized so that 100 pixels corresponded to 1 inch for all participants. After successfully resizing the screen, participants completed all three tasks. During the task, the cursor was hidden from the screen to minimize distractions. After completing the attention tasks, participants completed the questionnaire battery and demographic questions. Finally, we asked participants whether they ever got up or were interrupted during the study, and how noisy their environment was during the attention tasks. The full experiment took ~35 minutes. Participants were paid £4.38 when they reached the end of the experiment.

**Cognitive measures**

The attention tasks were programmed in JsPsych version 3.6.1 (De Leeuw, 2015). For all materials and links to working versions of the tasks, see the Github repository.

**Flanker task.** The Flanker task measures selective attention and response inhibition (B. A. Eriksen & Eriksen, 1974). The Flanker task began with eight practice trials, followed by 64 test trials. On each trial, participants saw a set of five arrows pointing either left or right. Participants were instructed to indicate the direction of the central arrow by pressing the respective arrow keys, while ignoring the flanking arrows to the left and right. All trials included black arrows against a white background. in the *congruent* trials (50%), the flanking arrows pointed in the same direction as the central arrow. in the *incongruent* trials (50%), the arrows pointed in the opposite direction. The arrows were randomly presented in the top-half or bottom-half of the screen. Each trial started with a fixation cross (1000 ms), after which the arrows were visible until a response was given. Participants received performance feedback during the practice trials, but not during the test block.

**Cued Attention task.** The Cued Attention task was an adapted version of the Posner task, which measures the speed of attention for peripheral cues (Posner, 1980). The Cued Attention task began with eight practice trials, followed by 64 test trials. On each trial, a left- or right-pointing arrow was presented in one of eight random locations at 300 pixels from the center of the screen. Participants were instructed to indicate the direction of the arrow by pressing either the left- or right arrow key on their keyboard. All trials included a black cue and arrow against a white background. On *cued* trials (50%), a cue ('*') preceded the arrow in the exact same location. On *neutral* trials (50%), the cue preceded the arrow, but appeared at the center of the screen (not where the arrow would appear). Thus, the cue was perfectly predictive of the target location on cued trials, but provided no predictive information about the location of the arrow on neutral trials. Each trial started with a fixation cross at the center of the screen for 1000 ms. Then, the cue appeared for 250 ms, followed by the target arrow, until a response was given.

**Change Detection task.** The Change Detection task measures the ability to detect subtle spatial changes. The Change Detection task started with five practice trials followed by 50 test trials. On each trial, participants saw five colored circles (red, lightblue, dark-blue, yellow, and purple) against a gray background, each with a radius of 15 pixels. Each circle was located in a semi-random location around the central fixation cross. The location of each circle was sampled within a pre-specified area of 50-by-50 pixels to prevent overlap. Participants had 1000 ms to memorize the locations of the five circles. Then, the circles disappeared for 500 ms and then reappeared. On *change* trials (50%), one of the circles had moved to another location with a fixed displacement of 40 pixels in a 360 degree direction. On *no change* trials (50%), all circles were still in the same location. Participants were instructed to indicate whether all circles were still in the same location or one of the circles had changed location by pressing the left- or right-arrow key. The displacement of *one* circle was the only potential difference on each trial;

**DDM/SSP parameters.** We analyzed Flanker task performance with the SSP model (Grange, 2016; White et al., 2011, 2018; White & Curl, 2018), using the *flankr* package (Grange, 2016). For each participant, the SSP provided us with estimates of: (1) strength of perceptual input (general quality of information that participants get from the arrows), (2) interference (initial attention width divided by the speed at which attention is narrowed down to the central arrow), (3) non-decision time (combination of speed of initial stimulus encoding and response execution), and (4) boundary separation (response caution). We always fixed the starting-point to the midpoint between the two boundaries, as modeling bias makes little sense when the boundaries correspond to correct and incorrect responses (as is the case here), rather than the distinct response options. Our focus on interference as a ratio between attention width and shrinking rate deviated from the preregistration, as we initially planned to investigate both aspects of attention separately. However, we discovered that both parameters in isolation were unreliable because of an inherent trade-off, while the ratio did provide a stable measure. This was supported in a simulation study by White et al. (2018) showing that the ratio measure is reliable. See the supplemental materials (section 3) for a comparison between the preregistered and the updated analyses.

For the Change Detection task and the Cued Attention task, we used a hierarchical Bayesian implementation of the standard DDM (HDDM). For each participant, the HDDM provided us with estimates of: (1) drift rate (speed of information accumulation; analogous to strength of perceptual input in the SSP model, except that drift rate is time-invariant), (2) non-decision time (same as in the SSP model), and (3) boundary separation (same as in the SSP model). The hierarchical Bayesian fitting procedure was a deviation from the preregistration, in which we planned to use Maximum Likelihood (ML) estimation. There were several issues with estimating DDM parameters for the Cued Attention task, which we later discovered were caused specifically by ML. An important difference between HDDM and ML is that HDDM uses the group information to inform individual parameter estimates, whereas ML models are fitted to each individual separately. The hierarchical approach generally improves generally improves the accuracy of the estimation. See the supplemental materials (section 3) for an overview of the fit procedure and model fit across all studies.

**Self-report measures**

See Table 4.2 for bivariate correlations between measures of adversity across all studies.

**Violence exposure.** We measured violence exposure using the Neighborhood Violence Scale (NVS) and two items assessing involvement in violence before age 13 (Frankenhuis, Vries, et al., 2020; Frankenhuis & Bijlstra, 2018; Young et al., 2022). The NVS contains seven items measuring perceived exposure to violence before age 13 (e.g., “Crime was common in the neighborhood where I grew up”). Participants rated each on a scale from 1 (never true) to 5 (very often true). The physical fighting items assessed the number of times participants witnessed fights before age 13: “Based on your experiences, how many times did you see or hear someone being beaten up in real life, before age 13?” and “How many times were you in a physical fight, before age 13?” Answers to both items ranged from 1 (0 times) to 8 (12 or more times). The items of the NVS were averaged together (Cronbach’s $\alpha$ = 0.92). Similarly, we averaged the scores on the two fighting items together. For the main analyses, we created a perceived violence exposure composite by standardizing the NVS and fighting composites and calculating an unweighted average.

**Environmental unpredictability.** We included five measures of environmental unpredictability across different temporal scales: (1) the Questionnaire of Unpredictability in Childhood (QUIC; Glynn et al., 2019); (2) the Perceived Childhood Unpredictability scale (Young et al., 2018); (3) the Confusion, Hubbub, and Order Scale (CHAOS; Matheny et al., 1995); (4) stability of the family and social environment; and (5) objective indicators of unpredictability. All scales were adapted to refer to experiences before age 13. We computed a composite measure of all z-transformed unpredictability measures. See section 2 of the supplemental materials for an exploration of the factor structure of these measures.

The QUIC captures environmental and household unpredictability. We made three preregistered changes to the original scale (Glynn et al., 2019), to better align it with the other scales. First, all items were rated on a scale of 1 (never true) to 5 (very often true), except for four items referring to specific experiences (e.g., "I experienced changes in my custody arrangement"). For these items, we adopted a response scale with the options "never", "only once", "a couple times", "several times", "many times". Second, quantifiers such as "frequently", "often", and "There was a period of time when […]" were dropped to better match the response scale. Third, we excluded the item "My parents got divorced" because it did not fit the new response labels and this information was already captured by one of the items of the perceived unpredictability scale. Reliability of the scale was high (Cronbach's $\alpha$ = 0.95).

The perceived childhood unpredictability scale included eight items measuring perceived unpredictability before age 13 (e.g., "My family life was generally inconsistent and unpredictable from day-to-day"). Participants rated each on a scale from 1 (never true) to 5 (very often true). Reliability of the scale was high (Cronbach's $\alpha$ = 0.91).

The CHAOS consists of 15 items measuring the level of chaos in the household (e.g., "No matter how hard we tried, we always seemed to be running late"). All items were rated on a scale of 1 (never true) to 5 (very often true) instead of the original yes/no answer format. Reliability of the scale was high (Cronbach's $\alpha$ = 0.93).

We included one additional scale to measure the stability of the family and social environment. On a scale of 1 (the same all the time) to 5 (constant and rapid changes), participants indicated how often the following aspects of their family and social environment changed before age 13: (1) economic status; (2) family environment; (3) childhood neighborhood environment; and (4) childhood school environment.

Finally, we included four objective measures of unpredictability before age 13: 1) "How often did you move?"; 2) "How many adults lived in your home on average?"; 3) "How many romantic partners did your mother have (not counting your father)?"; 4) "How many romantic partners did your father have (not counting your mother)?". Previous studies have found associations between (subsets of) these measures and subjective measures of adversity as well as with developmental outcomes (Belsky et al., 2012; Ellis et al., 2009; Young et al., 2022).

**Table 4.2.** Pooled bivariate correlations and descriptive statistics of measures of childhood violence exposure and environmental unpredictability across the three studies.

| | Violence exposure | | | Environmental unpredictability | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1. Neigh. violence | - | | | | | | | | | |
| 2. Fighting | 0.50*** | - | | | | | | | | |
| 3. Violence comp. | 0.87*** | 0.86*** | - | | | | | | | |
| 4. QUIC | 0.52*** | 0.46*** | 0.56*** | - | | | | | | |
| 5. Perc. unpredictability | 0.36*** | 0.32*** | 0.39*** | 0.81*** | - | | | | | |
| 6. CHAOS | 0.46*** | 0.41*** | 0.50*** | 0.84*** | 0.79*** | - | | | | |
| 7. Env. change | 0.36*** | 0.35*** | 0.43*** | 0.59*** | 0.50*** | 0.45*** | - | | | |
| 8. Obj. unpredictability | 0.39*** | 0.32*** | 0.37*** | 0.56*** | 0.56*** | 0.40*** | 0.73*** | - | | |
| 9. Subj. Unpredictability | 0.47*** | 0.44*** | 0.53*** | 0.93*** | 0.94*** | 0.94*** | 0.54*** | 0.51*** | - | |
| 10. Unpredictability comp. | 0.49*** | 0.45*** | 0.53*** | 0.89*** | 0.87*** | 0.82*** | 0.71*** | 0.81*** | 0.92*** | - |
| Mean | 1.94 | 1.97 | -0.01 | 2.13 | 2.11 | 2.41 | 1.83 | -0.01 | -0.02 | -0.01 |
| SD | 0.83 | 1.34 | 0.85 | 0.72 | 0.98 | 0.83 | 0.78 | 0.69 | 1.00 | 0.74 |
| Median | 1.71 | 1.50 | -0.28 | 2.03 | 1.88 | 2.33 | 1.75 | -0.21 | -0.20 | -0.17 |
| Min | 1.00 | 1.00 | -0.98 | 1.00 | 1.00 | 1.00 | 1.00 | -0.85 | -1.59 | -1.15 |
| Max | 5.00 | 8.00 | 3.99 | 4.84 | 5.00 | 4.87 | 5.00 | 5.37 | 3.46 | 3.97 |
| Skew | 1.36 | 2.03 | 1.48 | 0.65 | 0.81 | 0.40 | 1.35 | 2.35 | 0.63 | 1.08 |
| Kurtosis | 1.65 | 4.67 | 2.34 | 0.05 | -0.27 | -0.45 | 2.08 | 8.04 | -0.21 | 1.52 |

Note: * = $p$ < .05, ** = $p$ < .01, *** = $p$ < .001. CHAOS = Chaos, Hubbub, and Order Scale; Env. change = environmental change; Obj. unpredictability = objective unpredictability; Neigh. violence = neighborhood violence; Perc. unpredictability = perceived unpredictability; QUIC = Questionnaire of Unpredictability in Childhood; Subj. unpredictability = subjective unpredictability; SD = standard deviation; Unpredictability comp. = unpredictability composite; Violence comp = violence composite.

### Data analyses

**Multiverse analysis.** In an amendment to the preregistration, we quantified the robustness of our findings against six data cleaning decisions that may affect the robustness of online studies by using multiverse analysis, using the *multitool* package (Young & Vermeent, 2023). Multiverse analysis allows for systematically evaluating the robustness of analyses across all combinations of different arbitrary data processing decisions (for details, see Del Giudice & Gangestad, 2021; Simonsohn et al., 2020; Steegen et al., 2016). Specifically, we looked at the influence of including or excluding 1) participants who scored below 0.5 on a build-in bot-detection measure on Prolific (potentially indicating a bot); 2) participants who did not rescale their screen at the start of the experiment; 3) participants who did not enter fullscreen mode prior to starting the tasks; 4) participants who exited fullscreen mode at any point during the tasks; 5) participants who indicated high levels of noise in their environment; 6) participants who indicated extreme interruptions during the experiment. See the supplemental materials (section 5) for figures summarizing $p$-distributions and the explained variance in the regression coefficients of each data cleaning decision.

**Confirmatory analyses.** For the Cued Attention and Flanker task RTs, we used linear mixed effects models to test violence exposure x task condition (sum-coded) interactions on mean RTs (calculated separately for each condition) and each DDM parameter. All mixed effects models included a random intercept for participants. For the Change Detection and Flanker task SSP parameters, we used linear regression models to test the main effect of adversity on mean RTs and each DDM/SSP parameter. We did not analyze accuracy rates as these were close to ceiling for the Flanker and Cued Attention task. To meet model assumptions of normally distributed residuals, mean reaction time were log-transformed, separately for the congruent and incongruent condition. Analyses involving interference (Flanker task) and boundary separation (all tasks) parameters violated the assumption of normally distributed residuals. For boundary separation, we solved this using log-transformation. For interference, non-normality was caused by extreme outliers (>3.2SD), which we excluded from the analyses.

**Results and discussion**

Table 4.3 summarizes the results. In the flanker task, more violence exposure was associated with lower strength of perceptual input under 31.25% of multiverse specifications (although the median 95% CI interval contained zero). We additionally found a significant main effect of violence exposure on interference under 100.00% of multiverse specifications, such that more violence exposure was associated with less interference. This was contrary to our expectation that people exposed to adversity would have more difficulties dealing with interference from irrelevant distractors.

Participants with more exposure to childhood violence were slower in the Cued Attention task, which was mainly related to a higher level of response caution (boundary separation). in the Change Detection task, more childhood violence exposure was associated with slower speed of information processing (drift rate) under 50.00% of multiverse specifications, but not with longer RTs. These results were not in line with our expectation that people from adversity would perform better on cognitive tasks that require a broad, present-focused attention style.

Exploratory analyses did not show any significant associations with Flanker task performance. Participants with more exposure to childhood unpredictability were slower in the Cued Attention task (main effect) (median $\beta$ = 0.11, 95% CI = [0.02, 0.19], 81.25 % of *p*s <.05), which was related to slower non-decision time (median $\beta$ = 0.10, 95% CI = [0.02, 0.17], 100.00 % of *p*s <.05). We did not find a significant association between exposure to childhood unpredictability and mean RTs on the Change Detection task, although more unpredictability was negatively associated with drift rates (median $\beta$ = −0.10, 95% CI = [−0.20, −0.00], 53.12 % of *p*s <.05).

**Table 4.3.** Main and interaction effects of the effect of violence exposure on task performance.

| | Main Effect | | | Interaction | | |
|---|---|---|---|---|---|---|
| | β | 95% CI | *p* (%) | β | 95% CI | *p* (%) |
| **Cued Attention Task** | | | | | | |
| Raw response time | 0.10 | [0.01, 0.19] | 67.19 | 0.01 | [-0.01, 0.02] | 0 |
| Drift rate | 0.00 | [-0.08, 0.09] | 0.00 | -0.04 | [-0.08, 0.01] | 31.25 |
| Non-decision time | 0.05 | [-0.03, 0.13] | 0.00 | -0.02 | [-0.07, 0.02] | 9.375 |
| Boundary separation | 0.10 | [-0.01, 0.20] | 43.75 | | | |
| **Change Detection Task** | | | | | | |
| Raw response time | 0.05 | [-0.05, 0.15] | 0.00 | | | |
| Drift rate | -0.10 | [-0.20, 0.00] | 50.00 | | | |
| Non-decision time | -0.04 | [-0.14, 0.06] | 0.00 | | | |
| Boundary separation | 0.05 | [-0.05, 0.16] | 12.50 | | | |
| **Flanker Task** | | | | | | |
| Raw response time | 0.05 | [-0.04, 0.14] | 0.00 | -0.02 | [-0.04, -0.00] | 100 |
| Perceptual input | -0.08 | [-0.18, 0.02] | 31.25 | | | |
| Interference | -0.17 | [-0.26, -0.07] | 100.00 | | | |
| Non-decision time | 0.06 | [-0.04, 0.17] | 15.62 | | | |
| Boundary separation | -0.03 | [-0.13, 0.07] | 0.00 | | | |

*Note:* The p (%) column reflects the number of analyses that produced p-values < .05 for a given multiverse.

The pattern of findings in the Flanker task was interesting for two reasons. First, the Flanker task is a widely used task to assess the ability to inhibit irrelevant information, and people exposed to adversity typically show lowered performance. Our pilot results, though, suggest that lowered performance may not be caused by a reduced ability to inhibit distracting information. Instead, people exposed to adversity might have a lower strength of perceptual input, leading to slower and less efficient information processing. If true, these initial findings suggest that performance might be improved through interventions that increase the visual quality of stimuli. In Study 1, we aimed to replicate and extend these findings.

## 4.3 Study 1

The goal of Study 1 was to follow up on the Pilot study by manipulating the visual quality of information in the Flanker task. Participants completed three versions: a standard version (similar to the Pilot study), one with enhanced visual information, and one with degraded visual information. We again focused on childhood exposure to violence. Our first aim was to examine the robustness of our finding of improved interference control in the Flanker task in relation to more adversity exposure in the Pilot study. We did so by analyzing the data of the standard condition, as well as by pooling the data of the Pilot study and Study 1. Our second aim was to investigate whether manipulating visual information in the Flanker task would influence performance for people with more violence exposure.

We preregistered two potential data patterns and associated interpretations, without favoring one over the other *a priori*. First, the strength of perceptual input might be lower for people with more exposure to violence compared to people with less exposure to violence across all conditions. Second, lower performance in the standard version might reflect an adaptive trade-off towards cognitive functioning that is less affected by noise or perturbations, at a cost of lower overall performance (Del Giudice & Crespi, 2018). In that case, we would expect the strength of perceptual input to be influenced to a lesser extent across conditions for people with more exposure to violence than for people with less exposure to violence. As a result, they might not benefit as much from enhanced visual information, yet might be able to better maintain performance with degraded information.

## Methods

### Participants

Participant recruitment was identical to the Pilot study. In total, 567 people from the United States between the ages of 18 and 30 participated (See Table 4.1). We obtained ethical approval from the Ethics Review Board of the Faculty of Social & Behavioral Sciences of Utrecht University (FETC20-490). We applied the same exclusion criteria as reported in the Pilot study. The final sample consisted of 497 participants.

### Flanker task

We programmed the Flanker task in JsPsych version 6.3.1 (De Leeuw, 2015) with three conditions. Each condition consisted of eight practice trials, followed by 64 test trials. In the *standard* condition, the arrows were 40 pixels in size (0.4 inches) and had zero padding between them. In the *enhanced* condition, we increased the arrow size by 12.5% to 45 pixels (0.45 inches), and increased the space between the arrows to 5 pixels. This increased the width of the stimulus display by 50% with respect to the standard display. In the *degraded* condition, sizes and space between arrows were the same as in the standard version, but all arrows were rotated 45°. The lines of the arrows always had the same 45° angle. For example, if the flanking arrows pointed to the upper-left on an incongruent trial, the central arrow pointed to the lower-right. On congruent trials, all arrows pointed in the same direction (e.g., upper-right). Participants completed each condition separately in different blocks, in randomized order.

### Self-report measures

The self-report measures were identical to those used in the Pilot study.

### Procedure

The procedure was identical to the Pilot study. The full experiment took approximately 30 minutes. Participants were paid £3.75 after they completed the full study.

### Data analyses

**Multiverse analysis.** We included the same arbitrary decisions in the multiverse analyses as in the Pilot study. For the pooled analyses—i.e., joint analysis of the Pilot study and the standard condition of Study 1—there was one minor change in how we included screen rescaling as a preprocessing decision in the multiverse. In Study 1, we changed the screen rescaling procedure by converting the initial size of the resize box to 300 pixels instead of 100 pixels. This way, the stimulus display would still be close to the intended size if participants did not engage in any resizing. However, this led to one important change for the pooled analysis: rescaling (yes or no) was included as an arbitrary exclusion decision in the multiverse analyses with four combinations: (1) exclude non-scalers in both studies; (2) include non-scalers in both studies; (3) exclude non-scalers in the Pilot study, include non-scalers in Study 1; (4) include non-scalers in the Pilot study, exclude non-scalers in Study 1.

For each analysis, we report the median $\beta$s, 95% confidence intervals, and the proportion of $p$-values < .05 across all analytic decisions. For the confirmatory analyses, we used bootstrapping to compute the probability of obtaining an effect size at least as extreme as observed in the real data, conditioned on a true effect size of zero (for details, see Simonsohn et al., 2020). See the supplemental materials (section 5) for figures summarizing $p$-distributions and the explained variance in the regression coefficients of each data cleaning decision.

**Confirmatory analyses.** To address the first aim, we analyzed the data from the standard condition, as well as pooled the Flanker task data of the Pilot study and the current study. We ran separate linear models for each SSP parameter as well as RT difference scores (based on log-transformed mean RTs of each condition) with violence exposure as main predictor and study as covariate (effect-coded). To address the second aim, we analyzed the effect of violence exposure and Flanker task condition type on performance using linear mixed effects models with a random intercept per participant. The five main dependent variables were mean RT difference (based on log-transformed mean RTs of each condition) and the SSP parameters: Perceptual input, boundary separation, non-decision time, and interference. For each outcome measure, we ran two separate models: one comparing the standard condition with the enhanced condition, and one comparing the standard condition with the degraded condition. In both models, condition was dummy-coded using the standard condition as the reference group.

The use of RT difference scores differed from the Pilot study, where we included task condition as a moderator. We opted for RT difference scores here (as well as in Study 2) to prevent the use of three-way interactions, for which we did not have enough power.

## Results and discussion

### Standard Flanker performance

Table 4.4 summarizes the multiverse results for the effects of violence exposure (confirmatory analysis) and unpredictability (exploratory analysis). Unlike in the Pilot study, we did not find any significant associations with violence exposure. In the exploratory analysis, there was a significant negative association between unpredictability and perceptual input (median $\beta$ = −0.12, 95% CI = [−0.22, −0.03], 100.00 % of $p$s <.05).

**Table 4.4.** Standardized effects of violence exposure and unpredictability on Flanker performance in study 1.

| | β | 95% CI | *p* (%) | *p* |
|---|---|---|---|---|
| **Violence exposure (confirmatory)** | | | | |
| $RT_{difference}$ | 0.04 | [-0.06, 0.13] | 0.00 | .476 |
| Perceptual input | -0.02 | [-0.12, 0.07] | 0.00 | .596 |
| Interference | 0.02 | [-0.07, 0.12] | 0.00 | .630 |
| Non-decision time | -0.01 | [-0.10, 0.09] | 0.00 | .850 |
| Boundary separation | 0.03 | [-0.07, 0.12] | 0.00 | .542 |
| **Unpredictability (exploratory)** | | | | |
| $RT_{difference}$ | -0.03 | [-0.12, 0.07] | 0.00 | .596 |
| Perceptual input | -0.12 | [-0.22, -0.03] | 100.00 | .046 |
| Interference | 0.09 | [-0.01, 0.18] | 34.38 | .126 |
| Non-decision time | -0.03 | [-0.12, 0.07] | 0.00 | .592 |
| Boundary separation | -0.04 | [-0.14, 0.05] | 0.00 | .362 |

*Note:* The p (%) column reflects the number of analyses that produced p-values < .05 for a given multiverse. We computed overall p-values using a bootstrapped resampling method, which reflect the probability of obtaining an effect size as extreme or more extreme given the median effect is 0.

After pooling the data of the Pilot study and Study 1, there was a negative association between violence exposure and interference (median $\beta$ = −0.07, 95% CI = [−0.14, −0.00], 64.06 % of $p$s <.05, bootstrapped $p$ = .028). Violence exposure was associated with lower strength of perceptual input under 64.06% of multiverse specifications, but the bootstrapped $p$-value was not significant (median $\beta$ = −0.05, 95% CI = [−0.12, 0.01], bootstrapped $p$ = .100). We did not find other significant associations for either violence exposure or unpredictability.

### Flanker task conditions

The main effects of task condition on the strength of perceptual input were in the expected direction: relative to the standard condition, the quality of perceptual input was higher in the enhanced condition (median $\beta$ = 0.09, 95% CI = [0.04, 0.13], 100.00 % of $p$s <.05) and lower in the degraded condition (median $\beta$ = −0.13, 95% CI = [−0.18, −0.08], 100.00 % of $p$s <.05). Interference was lower in the enhanced condition (median $\beta$ = −0.26, 95% CI = [−0.31, −0.21], 100.00 % of $p$s <.05). Unexpectedly, interference was also lower in the degraded condition (median $\beta$ = −0.10, 95% CI = [−0.16, −0.04], 100.00 % of $p$s <.05), suggesting that the angle in the flanking arrows reduced interference, relative to the standard condition. However, none of the interaction effects for either violence exposure or unpredictability were significant (Table 4.5).

**Table 4.5.** Standardized interaction effects of violence exposure (confirmatory analysis) and unpredictability (secondary analysis) on Flanker performance across standard, enhanced, and degraded conditions.

| | Violence exposure X Condition | | | Unpredictability X Condition | | |
|---|---|---|---|---|---|---|
| | β | 95% CI | *p* (%) | β | 95% CI | *p* (%) |
| **Standard - Enhanced** | | | | | | |
| RT | -0.01 | [-0.06, 0.04] | 0.00 | -0.03 | [-0.09, 0.02] | 3.12 |
| Perceptual input | 0.03 | [-0.01, 0.08] | 0.00 | 0.02 | [-0.03, 0.07] | 0.00 |
| Interference | -0.01 | [-0.06, 0.04] | 0.00 | -0.04 | [-0.09, 0.01] | 25.00 |
| Non-decision time | 0.00 | [-0.05, 0.05] | 0.00 | 0.02 | [-0.03, 0.07] | 0.00 |
| Boundary separation | 0.01 | [-0.04, 0.06] | 0.00 | 0.00 | [-0.05, 0.06] | 0.00 |
| **Standard - Degraded** | | | | | | |
| RT | 0.03 | [-0.03, 0.09] | 6.25 | -0.01 | [-0.07, 0.05] | 0.00 |
| Perceptual input | 0.01 | [-0.04, 0.06] | 0.00 | 0.04 | [-0.02, 0.09] | 0.00 |
| Interference | 0.01 | [-0.05, 0.06] | 0.00 | -0.03 | [-0.09, 0.02] | 4.69 |
| Non-decision time | -0.02 | [-0.07, 0.04] | 0.00 | -0.02 | [-0.07, 0.04] | 0.00 |
| Boundary separation | 0.03 | [-0.03, 0.08] | 0.00 | 0.04 | [-0.02, 0.10] | 0.00 |

*Note:* Task conditions were dummy-coded with the standard condition as the reference. The p (%) column reflects the number of analyses that produced p-values < .05 for a given multiverse.

The results of the Pilot study and Study 1 were inconsistent with regard to the association between adversity and interference, but hinted at two general patterns. First, violence exposure was not associated with *increased* interference; instead, we found either the opposite effect or no effect. Second, both violence exposure and unpredictability were associated with lowered strength of perceptual input, albeit inconsistently. These findings, if replicable, are intriguing as they would suggest that the common finding of lowered Flanker task performance among people with more adversity exposure do not actually result from worse interference control—as typically inferred. Rather, such lowered performance would result from processes other than inhibition ability, such as slower general processing. Though interesting, our findings so far leave open the question why adversity might be negatively associated with perceptual input. This question was the focus of Study 2.

## 4.4 Study 2

Study 2 set out to compare two explanations for the finding that people exposed to adversity tended to show lower strength of perceptual input in the Flanker task. First, lowered strength of perceptual input in people exposed to adversity may indicate a difficulty in extracting relevant information (i.e., about their direction) from the arrows in general. Second, the difference in perceptual input may not be a cognitive deficit per se, but instead could be a signature of a difference in processing style—that is, a feature, and not a bug. People exposed to adversity may process information more holistically, focusing more on the configuration of pieces of information rather than individual pieces of information. in the Flanker task, this would lower the depth of perceptual processing of any individual stimulus, thus resulting in lowered strength of perceptual input, as we observed in the Pilot study and Study 1.

We preregistered three aims focusing both on violence exposure and unpredictability. First, we expected to replicate our earlier findings that adversity was associated with lowered perceptual input and lower interference in people exposed to adversity. Second, we included a Global-Local task to investigate the hypothesis—based on the findings of the Pilot study and Study 1—that people with more adverse experiences would develop a more holistic style of information processing. Third, we planned to conduct a within-subjects analysis of Flanker and Global-Local task performance to assess whether people with lowered perceptual input in the Flanker task would also show a more global processing style (rather than a local processing style) in the Global-Local task.

## Methods

### Participants

Participants were 600 people from the United States between the ages of 18 and 30. Recruitment was identical to Study 1. We obtained ethical approval from the Ethics Review Board of the Faculty of Social & Behavioral Sciences of Utrecht University (FETC20-490). We conducted a simulation-based power analysis for the planned linear mixed models with the Global Local task (see GitHub). We determined that power of > .80 for a standardized interaction effect of 0.06, with sigma (noise) set to 0.7 (comparable to observed sigmas in the first two studies) would require 550 participants. We recruited 600 participants, with the expectation to have a final sample of 550 participants after exclusions. We applied the same exclusion criteria as reported in the Pilot study and Study 1. The final sample consisted of 551 participants.

### Measures

The measures of childhood violence exposure and environmental unpredictability were identical to Study 1. The Flanker task was identical to the standard version used in Study 1.

*Global-Local task*. The Global-Local task is a measure of global-local processing (Navon, 1977). Many different versions of this task exist in the literature. One key dimension on which they differ is whether the task measures focused attention (by cueing attention towards the global or local level prior to stimulus presentation) or divided attention (by having participants search for a target on both levels) (Lee et al., 2023). Here, we use a version measuring divided attention, which allows measuring whether someone tends to have a more global versus local processing style (Hakim et al., 2017; Lee et al., 2023; McKone et al., 2010).

Participants saw images of big, black letters (the global level) comprising small letters (the local level)—so-called Navon images (Navon, 1977)—against a white background. Participants first completed eight practice trials, after which they completed an additional 64 test trials. On each trial, participants searched for one of two target letters—an 'E' or 'H'—and indicated whether it was present on the global or local level by pressing 'g' or 'l' on their keyboard, respectively. Each stimulus was 600 pixels high and 395 pixels wide and comprised seven local letters vertically and five local letters horizontally. The stimuli consisted of combinations of the letters 'T', 'F', 'P', 'L', 'H', and 'E'. All stimuli always contained one (and only one) of the target letters 'H' and 'E' on either the local or global level. The other letters were randomly varied, and the global and local level never contained the same letter. Thus, the global-local task did not contain a congruent and incongruent condition as did the Flanker task.

**Procedure**

The procedure was identical to Study 1. The full experiment took ~30 minutes. Participants were paid £4.50 when they reached the end of the experiment.

**Data analyses**

**Multiverse analysis.** We included the same decisions in the multiverse as in the previous studies. However, there was one deviation from the preregistration: the multiverse analysis contained the same arbitrary decisions as the Pilot study and Study 1, instead of a subset, as we preregistered (for details, https://stefanvermeent.github.io/attention_project/preregistrations/README.html). See the supplemental materials (section 5) for figures summarizing *p*-distributions and the explained variance in the regression coefficients of each data cleaning decision.

**DDM estimation.** For the Flanker task, we used the SSP (Grange, 2016; White et al., 2011, 2018; White & Curl, 2018) using the same procedure as in Study 1. For the Global-Local task, we used a hierarchical Bayesian DDM to fit the data using the *runjags* package (Denwood, 2016). See the supplemental materials (Section 3) for more information about the procedure and model fit.

We deviated from our preregistration regarding the preprocessing of Global-Local task data. Specifically, we relaxed the low performance threshold as the task was more difficult than anticipated. These deviations are described in the supplemental materials (section 4).

**Confirmatory analyses.** We ran simple regressions for analyses involving only main effects (aim 1), and linear mixed effects models for analyses involving within-subject interactions (aim 2 and 3). To address aim 3 (within-subject interaction between Global-Local task drift rate and Flanker task strength of perceptual input), we further preprocessed the data in two steps. First, we computed a difference score of Global-Local drift rates by subtracting the drift rate on local trials from the drift rate on global trials (with higher scores reflecting relatively faster information processing on global trials). Second, we separately standardized the Flanker task strength of perceptual input and Global-Local task drift rate difference. We fitted linear mixed effects models with the standardized performance measures as the dependent variable, and adversity type, task (Flanker task or Global-Local task, sum-coded) and their interaction as independent variables.

**Results and discussion**

Figure 4.2 and 4.3 summarize the multiverse results for the effects of violence exposure and unpredictability within Study 2 and pooled across all studies. In Study 2, violence exposure was negatively associated with strength of perceptual input ($\beta_{median}$ = −0.18, 95% CI = [−0.26, −0.09], 100.00 % of *p*s < .05, bootstrapped *p* < .001), but not associated with interference ($\beta_{median}$ = −0.04, 95% CI = [−0.14, 0.05], 0.00 % of *p*s < .05, bootstrapped *p* = .672). Unpredictability was not associated with either strength of perceptual input ($\beta_{median}$ = −0.05, 95% CI = [−0.15, 0.04], 3.12 % of *p*s < .05, bootstrapped *p* = .026), nor with interference ($\beta_{median}$ = 0.03, 95% CI = [−0.06, 0.12], 0.00 % of *p*s < .05, bootstrapped *p* = .142).

In the pooled analysis, the results were similar for both types of adversity. Violence exposure was associated with lower strength of perceptual input ($\beta_{median}$ = −0.10, 95% CI = [−0.17, −0.04], 100.00 % of *p*s < .05, bootstrapped *p* < .001), but not with interference ($\beta_{median}$ = −0.01, 95% CI = [−0.08, 0.05], 0.00 % of *p*s < .05, bootstrapped *p* = .672). Similarly, unpredictability was associated with a lower quality of perceptual input ($\beta_{median}$ = −0.08, 95% CI = [−0.15, −0.02], 95.31 % of *p*s < .05, bootstrapped *p* = .026), but not with interference ($\beta_{median}$ = 0.05, 95% CI = [−0.01, 0.12], 0.00 % of *p*s < .05, bootstrapped *p* = .142).

**Global-Local task performance**

There was a main effect of violence exposure on Global-Local drift rates, with more violence exposure being associated with slower speed of information processing ($\beta_{median}$ = −0.19, 100.00% of *p*s < .05, bootstrapped *p* < .001). There also was a main effect of task condition on drift rates, with people processing information faster when the target was present at the global level compared to the local level, ($\beta_{median}$ = 0.05, 100.00% of *p*s < .05, bootstrapped *p* < .001). Finally, there was an interaction effect between violence exposure and task condition ($\beta_{median}$ = 0.04, 95.31 % of *p*s < .05, bootstrapped *p* = .038). Simple slopes analyses revealed that participants with lower levels of violence exposure did not differ in speed of processing of global versus local targets ($b_{median}$ = 0.01, 0.00% of *p*s < .05). In contrast, participants with higher levels of violence exposure processed global targets faster than local targets ($b_{median}$ = 0.08, 100.00% of *p*s < .05).

There was a significant main effect of unpredictability on drift rates, with more unpredictability being associated with slower speed of information processing, ($\beta_{median}$ = −0.10, 95% CI = [−0.20, −0.01], 62.50 % of *p*s < .05, bootstrapped *p* = .024). We also found a main effect of task condition on drift rates, with people processing information faster when the target was present at the global level compared to the local level, ($\beta_{median}$ = 0.05, 95% CI = [0.02, 0.08], 100.00 % of *p*s < .05, bootstrapped *p* < .001). We did

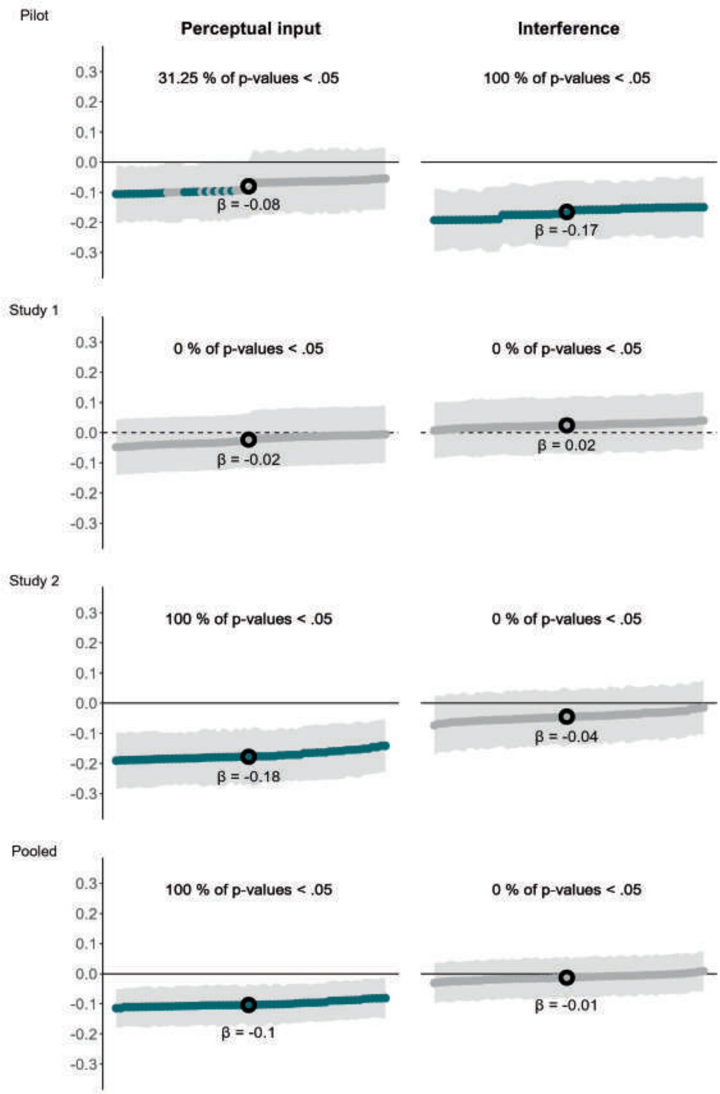



**Figure 4.2.** Multiverse results for the association between violence exposure with the strength of perceptual input and interference in the Flanker across all studies. Each panel depicts sorted beta coefficients across all combinations of arbitrary decisions (i.e., the effect curve across the whole multiverse). The top row depicts effect curves in the Pilot study. The second row depicts effect curves in Study 1. The third row depicts effect curves in Study 2. The fourth row depicts effect curves of the pooled analyses across all studies.

not find a significant unpredictability x task condition interaction effect ($\beta_{median}$ = 0.03, 95% CI = [−0.00, 0.06], 37.50 % of *ps* < .05, bootstrapped *p* = .100).

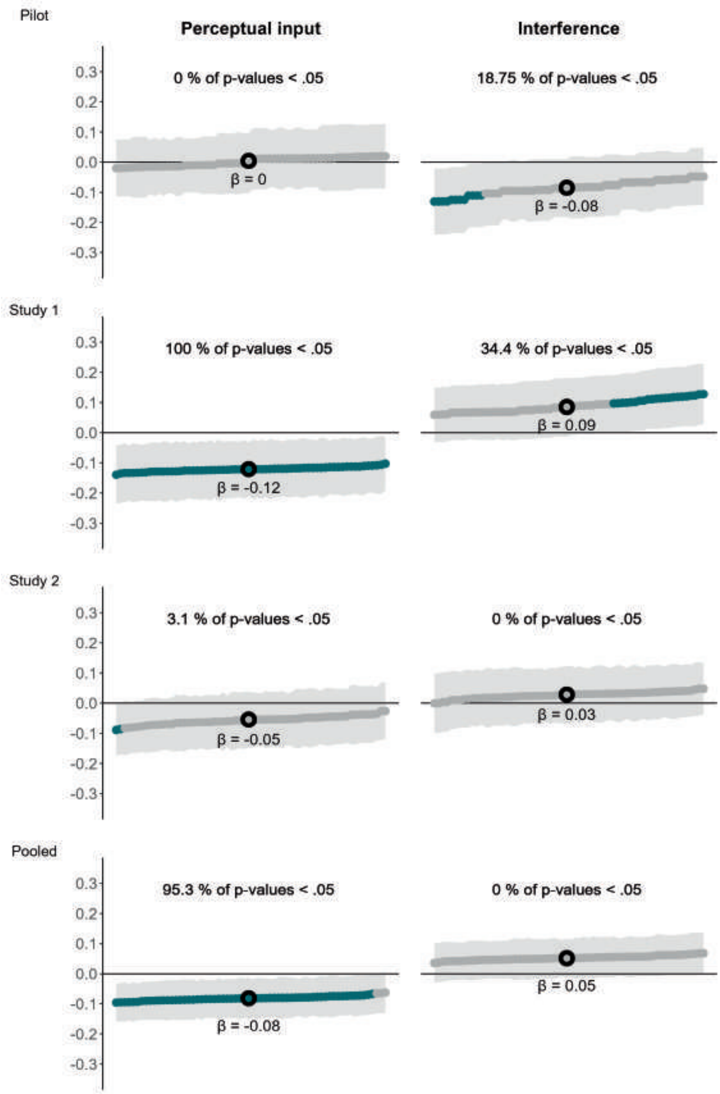



**Figure 4.3.** Multiverse results for the association between unpredictability with the strength of perceptual input and interference in the Flanker task across all studies. Each panel depicts sorted beta coefficients across all combinations of arbitrary decisions (i.e., the effect curve across the whole multiverse). The top row depicts effect curves in the Pilot study. The second row depicts effect curves in Study 1. The third row depicts effect curves in Study 2. The fourth row depicts effect curves of the pooled analyses across all studies.

**Within-subjects comparison of Flanker and Global-Local task information processing**

There were no significant main effects for violence exposure (bootstrapped $p$ = .486) nor for cognitive task (bootstrapped $p$ = .486). There was a significant interaction effect ($\beta_{median}$ = 0.15, 95% CI = [0.08, 0.21], 100.00 % of $p$s < .05, bootstrapped $p$ < .001) (See Figure 4.4). A simple slopes analysis revealed that people with higher levels of violence exposure showed lower strength of perceptual input in the Flanker task (b = −0.17, 100.00% of $p$s < .05), and showed a more global versus local processing style in the Global-Local task (b = 0.13, 92.19% of $p$s < .05).

There was no significant main effect for unpredictability (bootstrapped $p$ = .414). However, there was a significant interaction effect ($\beta_{median}$ = 0.08, 95% CI = [0.01, 0.14], 67.19 % of $p$s < .05, bootstrapped $p$ = .044). A simple slopes analysis revealed that people with higher levels of unpredictability did not differ in their strength of perceptual input in the Flanker task (b = −0.05, 3.12 % of $p$s < .05), but showed a more global versus local processing style in the Global-Local task (b = 0.09, 34.38% of $p$s < .05).

To sum up, Study 2 provided additional support for the basic finding that violence exposure and unpredictability were associated with lower strength of perceptual input but not with differences in interference; with the caveat that the associations for unpredictability only showed up in pooled analyses. People with more exposure to violence and unpredictability also processed information more slowly in the Global-Local task. In line with our expectations, childhood exposure to violence was associated with both lowered strength of perceptual input and a more holistic processing style. The same processing style was observed for participants with more exposure to unpredictability, although they did not show lowered strength of perceptual input.

## 4.5 Exploratory analyses

We hypothesized that the potential adaptive benefits of a more diffuse scope of attention in adverse conditions might be linked to the notion of a *present-oriented attention style* (Frankenhuis et al., 2016; Van Gelder & Frankenhuis, 2024). People with a present-oriented attention style (versus a more future-oriented attention style) are more geared towards processing information that is relevant for solving challenges and obtaining rewards in the here-and-now. A general tendency to be more attuned to the present (while disregarding the future) is sometimes referred to as a short-term mindset (Kübel et al., 2023; Van Gelder & Frankenhuis, 2024), which also includes tendencies to be more impulsive, to more steeply discount future rewards, and to be more sensation-seeking. Although short-term mindsets are associated with exposure to adversity (Ganschow et al., 2023), it is unclear how they are associated with performance on attention tasks. We explored bivariate correlations pooled across all studies between two indicators of short-term mindsets (impulsivity and future orientation) and SSP para-

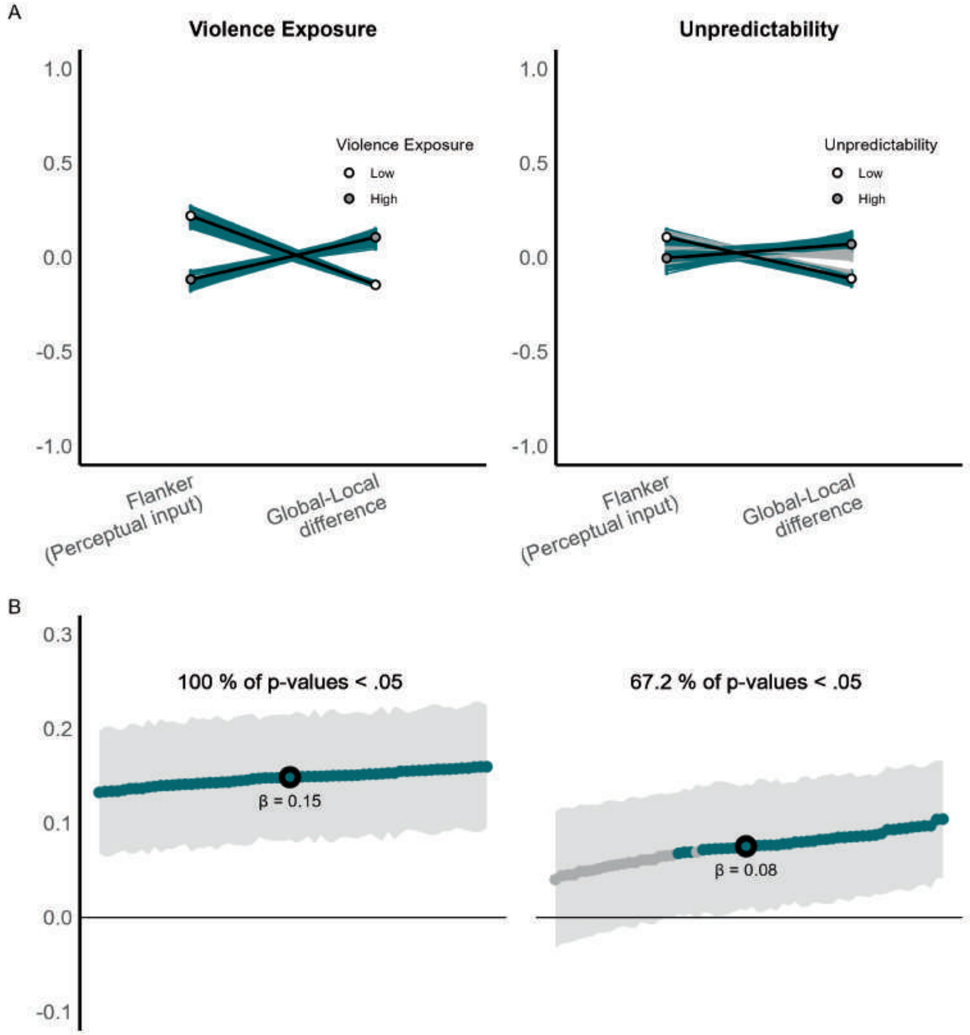



**Figure 4.4.** Multiverse results for the within-subjects comparison of Flanker and Global-Local task information processing. Panel A depicts the multiverse interaction effects, with the thick black lines denoting the median slope and the thin lines denoting effects for each combination of arbitrary decisions. Blue thin lines indicate significant effects ($p$ > .05), and grey thin lines indicate non-significant effects ($p$ > .05). Panel B depicts sorted beta coefficients across all combinations of arbitrary decisions (i.e., the effect curve across the whole multiverse). See the main text for more information about the multiverse analyses.

meters of the Flanker task (for more information on the measures, see section 1 of the supplemental materials).

See Table A2.2 for an overview of the correlations. Impulsivity was negatively associated with the strength of perceptual input ($r$ = -.07, $p$ = .004) and positively associated with interference ($r$ = .09, $p$ = .005). In addition, impulsivity was also associated with a more holistic information processing style ($r$ = .11, $p$ = .020). Thus, more impulsive participants processed information less deeply and were more easily drawn to distractions, but this might partly be explained by a holistic information processing style. Similarly, future-orientation was positively associated with perceptual input ($r$ = .09 $p$ < .001)—but not with interference ($p$ = .112)—in the Flanker task, and was also associated with a more detail-oriented processing style ($r$ = -.12, $p$ = .011). Thus, more future-oriented participants processed information more deeply, which might partly be explained by a detail-oriented processing style.

## 4.6 General discussion

We investigated how two dimensions of childhood adversity—violence exposure and environmental unpredictability—are related to differences in how people attend to and process information. Specifically, we hypothesized that exposure to adversity might lead to a present-oriented attention style that would facilitate rapidly detecting novel or changing information, yet which would interfere with ignoring distractors. Across one Pilot study and two main studies, we tested how adversity was associated with performance on different attention tasks. The Pilot study compared performance on a Cued Attention task, Change Detection task, and a Flanker task. Two follow-up studies focused in the Flanker task, with Study 2 also including a Global-Local task. We leveraged DDM to estimate the processes underlying lowered and improved performance. This allowed us to investigate whether performance differences were associated with abilities that are typically the main focus when using these tasks (i.e., attention orientation, interference control, information accumulation), or with other processes (e.g., stimulus encoding, response execution, response caution). Across all confirmatory and exploratory analyses, we leveraged multiverse analysis to systematically assess the robustness of our findings against several uncontrollable aspects of the online assessment (e.g., distractions, fullscreen exits).

### Main insights

We found little to no support for the presence of a present-oriented attention style in people exposed to adversity. More childhood exposure to violence was associated with slower processing of subtle changes in the Change Detection task and lower quality of perceptual input in the Flanker task. It was not associated with speed of processing of peripheral information in the Cued Attention task. Zooming in in the Flanker task, our two main studies found mixed evidence for the hypotheses that violence exposure and unpredictability were associated with lower strength of perceptual input. This mixed evidence suggests that people with more exposure to these adversities process information less deeply, leading to slower responses on congruent and incongruent trials in equal measure. This was corroborated by the pooled analyses across studies, which found that both exposure to violence and unpredictability were associated with lower strength of perceptual input, but not with differences in the ability to inhibit distractors. This finding contradicts the standard deficit interpretation of lowered performance on inhibition tasks by people with more adversity exposure (discussed below). In addition, lowered strength of perceptual input was associated with a holistic processing style. Thus, we did not find evidence that people with more violence exposure have more difficulties with inhibiting task-irrelevant information.

Our findings of the DDM decomposition of Flanker task performance challenge previous interpretations based on raw performance. Previous studies have found that people exposed to adversity and/or low SES backgrounds have longer RTs on incongruent trials relative to congruent trials (Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015; Noble et al., 2005), which is commonly interpreted as an impaired ability to inhibit irrelevant information. This fits with adaptive hypotheses, as inhibition is assumed to be useful mostly in stable and predictable environment that afford long-term goal pursuit, but can be costly in unpredictable and potentially dangerous environments (Daly & Wilson, 2005; Fields et al., 2021; Mittal et al., 2015). Contrary to previous studies, we found little to no evidence for performance differences on the basis of raw RTs. In addition, our DDM analyses showed that performance differences in the Flanker task are not driven by differences in interference control, but by more basic processes that are not typically considered when interpreting Flanker task performance. Although we are not aware of similar findings in the literature on adversity, comparable conclusions have recently been drawn in research on cognitive functioning related to depression and autism (Grange & Rydon-Grange, 2022; Merkt et al., 2013; Poole et al., 2024).

Our findings align with a broader literature that is critical of the validity of the Flanker task in particular, and that of cognitive control tasks more generally. As noted, several studies have failed to find coherent correlations between raw performance on different cognitive control tasks (e.g., Löffler et al., 2024; Rey-Mermet et al., 2019; Rouder & Haaf, 2019; Stahl et al., 2014). For example, previous research comparing several cognitive control tasks across different data sets using cognitive modeling found that shared variance between these tasks was mostly associated with processing speed and strategies (e.g., speed-accuracy trade-offs) (Hedge et al., 2022). Moreover, the modeling parameters reflecting conflict processing (similar to interference in our study) were barely correlated. Similarly, previous work has shown that individual differences on common EF tasks—among which the Flanker task—can be fully accounted for by general processing speed (Löffler et al., 2024). This literature, together with the findings reported here, underscore that researchers should be cautious when drawing inferences about cognitive control abilities in people exposed to adversity based on raw RTs and performance on individual tasks.

Finally, we showed that people exposed to adversity had a more holistic processing style, and that this style was associated with lower strength of perceptual input in the Flanker task. This could mean that people with more adversity exposure processed the Flanker task display more holistically; that is, focused less on individual arrows and more on the collection of arrows as a whole. One (tentative) interpretation is that in the absence of threatening or otherwise salient information, people with more exposure to adversity attend to and process information in the environment globally and less deeply. They might only shift to local processing of a single source of information if it seems threatening or otherwise salient (Schwabe et al., 2013; Shields et al., 2015). This would be consistent with research showing that growing up in a disadvantaged environment decreases the efficiency of the brain's (resting-state) salience network, which is in turn associated with lower raw performance on certain cognitive tasks (Cermakova et al., 2023; Gellci et al., 2019; Hilger et al., 2017; Yuan et al., 2012). This research also shows that in situations of acute stress, mental resources are reallocated to this salience network, increasing vigilance and facilitating adequate responding. Indeed, a few studies show that cognitive abilities that may be particularly relevant in adverse contexts—such as attention shifting and working memory updating—may be enhanced in people from adversity when they experience acute stress (Mittal et al., 2015; Young et al., 2018).

We did not control for (potentially) confounding variables in our models, even though variables like education, intelligence, and current adversity exposure generally correlate with both childhood adversity exposure and performance on the Flanker task. Our reason for not including them as covariates was that all these factors can be reasonably seen as mediators of the association between childhood adversity and cognitive performance. However, they are unlikely (or even impossible) causes of childhood adversity. Therefore, adjusting for these variables could have introduced bias to our estimation of the total effect of childhood adversity on performance (which was our estimand) (Rohrer, 2018). That said, one way in which our analyses may have been confounded is by using retrospective measures of adversity. For example, some work suggests that current psychopathology may bias retrospective reports of childhood adversity, although the causal pathways are still mostly unclear (Francis et al., 2023; Goltermann et al., 2023; Nivison et al., 2021; Patten et al., 2015). Systematic investigations into potential confounders will ultimately improve our understanding of the effects of early adversity (Ning et al., 2023).

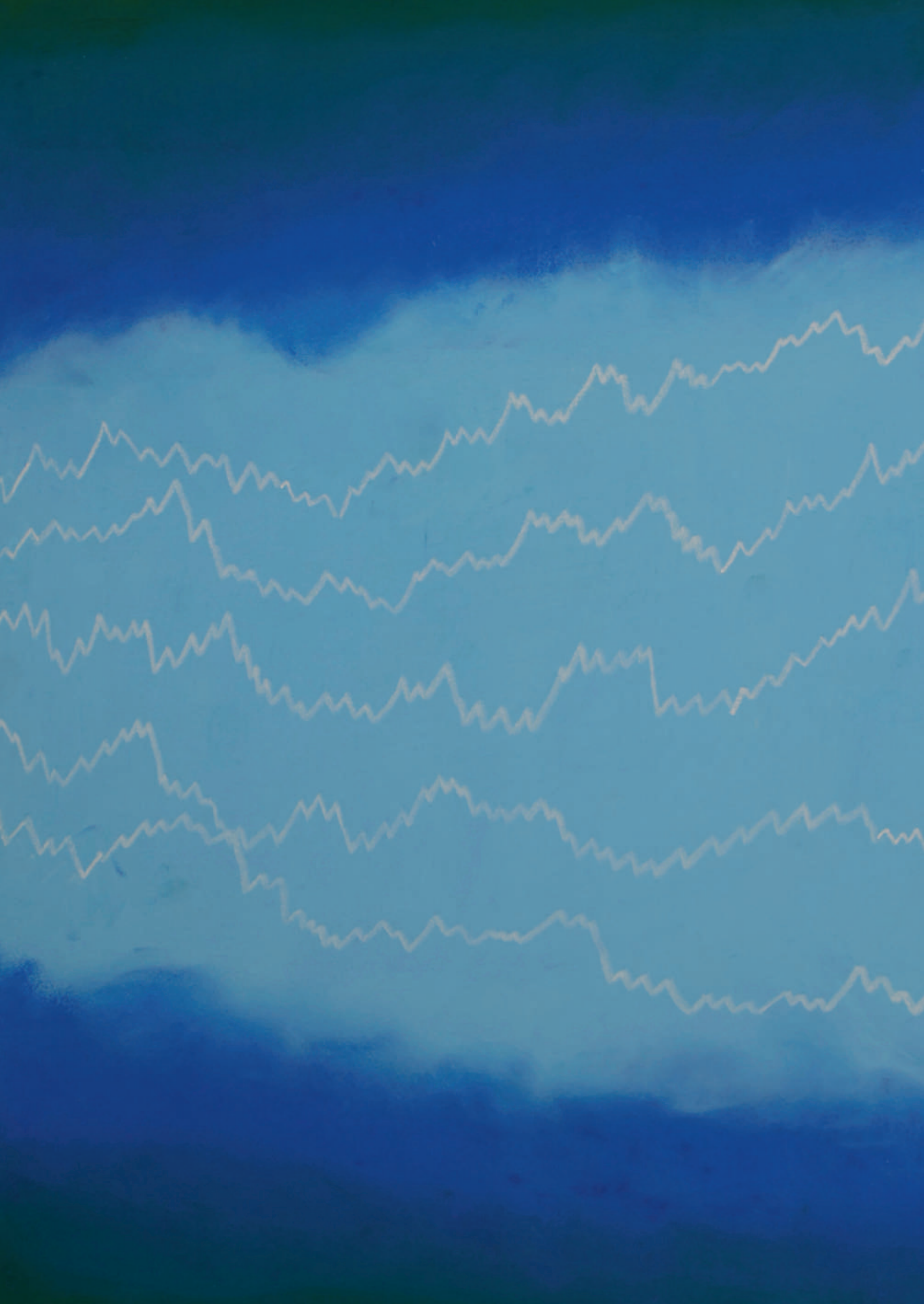
### Strengths, limitations, and future directions

Our study has three main strengths. First, each study included socioeconomically diverse participants. Second, the DDM allowed us to decompose performance in a more nuanced way than is possible with (typically used) raw performance scores. Third, the multiverse analyses provided a systematic overview of the robustness of our findings under different analytical decisions. Our study has three main limitations. First, all experiments were conducted online, which reduced control over people's testing environment, equipment, and behavior. Indeed, our results were the least robust against participants who skipped the screen-scaling procedure (to ensure the stimuli were adequately sized) and interruptions during the tasks, which are factors that are largely out of our control. Second, we deviated from our preregistrations in several ways in all studies, due to progressive insight. This decreased the severity of our statistical tests, and so this work would benefit from preregistered replications (Lakens, 2024).

Our findings suggest two main directions for future research. First, future studies could replicate and expand upon our finding that lower quality of information processing in people exposed to adversity is associated with a more holistic processing style. For example, future work could investigate whether people with more adversity exposure shift from holistic to a detail-oriented processing in situations of acute stress or otherwise salient information. Second, our results suggest that lower strength of perceptual input is likely the result of both processing styles, as well as of slower general processing. Future research could try to tease apart these sources using a within-subjects design simultaneously measuring inhibition, processing styles, and basic processing speed. Third, some research suggests that inhibition is not a unitary construct, instead distinguishing between response inhibition (which involves suppressing a prepotent response) and cognitive inhibition (which involves selective attention in the presence of distractors). Exposure to adversity might shape these two types of inhibition in different ways. For example, acute stress might impair performance on tasks of cognitive inhibition (of which the Flanker task is an example) and enhances performance on tasks of response inhibition (for a meta-analysis, see Shields et al., 2016; but see Dang, 2017). Future work could assess inhibition more broadly, e.g., by including tasks that are hypothesized to require cognitive or response inhibition.

## 4.7 General conclusion

We found that people with more childhood adversity exposure perform worse in the Flanker task not because of an impairment in their ability to inhibit distracting information, but because of lower strength of perceptual input. Our results suggest that people with more adversity exposure are not worse at inhibiting distractions; rather, they do not seem to process information in the environment deeply unless it proves to be a reliable and important source of information. These findings challenge dominant interpretations, which infer an inhibition deficit from lowered performance. This is an important difference not just for theory development, but also for future interventions aimed at closing performance gaps. For example, when applied to school contexts, interventions based on an inhibition interpretation would focus on the learning environment, perhaps removing things from the classroom that could be distracting. In contrast, an intervention based on an information processing interpretation might instead focus on increasing the apparent relevance of the learning materials, perhaps by providing more repetition or by making the content more ecologically relevant (Young et al., 2022). Thus, cognitive modeling can offer crucial insights for our understanding of cognitive abilities in adverse conditions.



# Chapter 5

## Inconclusive evidence for associations between adverse experiences in adulthood and working memory performance

**This chapter is based on**

Vermeent, S., Schubert, A.-L., DeJoseph, M.L., Denissen, J.J.A., van Gelder, J.-L., & Frankenhuis, W.E. (2025). Inconclusive evidence for associations between adverse experiences in adulthood and working memory performance. *Royal Society Open Science, 12*(1), 241837. https://doi.org/10.1098/rsos.241837

## 5.0 Abstract

Decades of research show that adversity tends to be associated with lower working memory (WM) performance. This literature has mainly focused on impairments in the capacity to hold information available in WM for further processing. However, recent adaptation-based studies suggest that certain types of adversity can leave intact, or even enhance, the ability to update information in WM. One challenge is that WM capacity and updating tasks tend to covary. Estimating the associations between adversity and different processes in WM requires isolating variance in performance related to WM capacity from variance in performance related to updating ability. In this Registered Report, participants from the Dutch Longitudinal Internet studies for the Social Sciences (LISS) panel completed two tasks measuring WM capacity and one task measuring both binding and updating of information. We measured participants' exposure to neighborhood threat, deprivation, and unpredictability. We estimated associations between adversity and latent estimates of WM capacity and updating using structural equation modeling. We did not find associations between adversity and WM capacity or updating, nor did we find evidence that the associations were practically equivalent to zero. Our results show that adversity researchers should account for overlap in WM tasks when estimating specific WM abilities.

**Acknowledgements**
We would like to thank Kathrin Sadus for advising us on the implementation of the working memory tasks.

**Author contributions**
All authors were involved in conceptualizing the study. SV coordinated the data collection and analyzed the data, and wrote the first draft of the manuscript. All authors provided feedback on the manuscript.

## 5.1 Introduction

Living in adverse conditions, with prolonged exposure to intense stress, tends to have a profound and enduring impact on cognitive functioning (Farah et al., 2006; Sheridan et al., 2022; Sheridan & McLaughlin, 2014). Although adversity can be described in many ways, we follow contemporary models focusing on threat, deprivation, and unpredictability as key dimensions of adversity (Ellis et al., 2009, 2022; McLaughlin et al., 2021; McLaughlin & Sheridan, 2016). A domain that seems to be particularly affected by adversity is working memory (WM). WM is a system for mentally building, maintaining, and updating immediately relevant information (Oberauer et al., 2018). Performance on WM tasks is associated with a host of social and cognitive abilities, such as math (Peng & Fuchs, 2016), reading (Chiappe et al., 2000), learning (Cowan, 2014), general intelligence (Conway et al., 2003), and mentalizing (Mutter et al., 2006). Not surprisingly, then, deficits in WM have negative consequences for both educational and professional outcomes (Ahmed et al., 2018; Alloway & Alloway, 2010; Guo et al., 2020; Spiegel et al., 2021). Decades of research show that adversity is generally negatively associated with performance on WM tasks (Goodman et al., 2019). However, emerging evidence suggests that specific aspects of WM might remain intact or even be enhanced through developmental adaptations to adversity. So far, the literature has tended to focus on related, but different aspects of WM in isolation, limiting a fuller integration. Here, we take a psychometric modeling approach to simultaneously examine potential decreases and enhancements in two WM components: capacity and updating.

### Deficit-based and adaptation-based models

A large literature has shown negative associations between exposures to adversity and performance on WM tasks (Farah et al., 2006; Sheridan et al., 2022; Sheridan & McLaughlin, 2014). These associations may be potentially attributable to the enduring influence of stress on several key brain regions that support WM (Duval et al., 2017; Hanson et al., 2012). Much of this work has focused on WM capacity, or the ability to keep multiple pieces of information simultaneously available for further processing. For early-life adversity, this negative association is already present during childhood, and persists into adulthood (Bos et al., 2009; Evans & Schamberg, 2009; Farah et al., 2006; Goodman et al., 2019; Hackman et al., 2010; Noble et al., 2007; but see Nweze et al., 2021). Studies with college students have found a link between both recent and lifetime experiences of stressful major life events (discrete negative events that have a clear onset and offset, unlike chronic adversity) with lower WM capacity (Klein & Boals, 2001; Shields et al., 2017, 2019).

The most common tasks used to examine the negative association between adversity and WM are simple span tasks (repeating a string of stimuli of increasing length), complex span tasks (remembering a string of stimuli while being engaged by a secondary task), and *n*-back tasks (judging whether the current stimulus in a string is identical to the stimulus *n* steps ago) (Goodman et al., 2019). Performance on these tasks is assessed through the number of items that participants can retain in WM, that is, their overall capacity (with the exception of *n*-back; for concerns about the construct validity of this task, see Frost et al., 2021; Kane et al., 2007).

Although both early-life and recent adversity appear to be negatively associated with WM capacity, a small set of studies suggest that exposure to adversity may leave intact, or even enhance, the ability to update items in WM in adolescents (Young et al., 2022) and adults (Young et al., 2018). Updating is defined as the ability to rapidly replace old information in WM with new information. The finding that updating may be left intact or even enhanced after exposure to adversity exemplifies emerging theoretical frameworks grounded in adaptive reasoning that are complementary to deficit frameworks (Ellis et al., 2017, 2022; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013).

Adaptation-based theories assume that developmental processes tailor an individual's cognitive abilities to the unique challenges and opportunities posed by their environment. The link between adversity and cognitive abilities is further assumed to be specific; as different types of adversity (e.g., threat vs. deprivation) pose different challenges, they should (at least in part) shape cognitive abilities in different ways. For example, with regards to executive functioning, some previous studies have found that children and adults with more exposure to unpredictability (characterized by random variation in adversity exposure over space or time) and threat tend to be better at rapidly shifting their attention between tasks (Fields et al., 2021; Mittal et al., 2015; Steudte-Schmiedgen et al., 2014; Young et al., 2022; but see Nweze et al., 2021). WM updating may be especially adaptive in unpredictable environments. WM updating allows people to maintain an up-to-date overview of the (changing) current state of the environment (Young et al., 2018). Additionally, improved WM updating performance has also been documented for threat exposure (Young et al., 2022). An enhanced WM updating ability could facilitate keeping track of and integrating signals that may potentially signal acutely threatening situations.

### Associations between WM capacity and updating

With deficit theories focusing on WM capacity and adaptation-based theories on WM updating, we may wonder how capacity and updating are related to each other. Performance on tasks measuring WM capacity and updating tend to be substantially correlated (in the range of .20-.50; Frischkorn et al., 2022; Löffler et al., 2024). This overlap appears to stem from shared demands of both types of tasks, in particular the need to create and maintain arbitrary bindings (Gruszka & Nęcka, 2017; Oberauer, 2009; Wilhelm et al., 2013). The term *binding* refers to the process of mapping memory items to specific positions in WM (e.g., serial, spatial, or temporal positions, depending on the task) (Oberauer, 2009; Oberauer & Lewandowsky, 2019). For example, on most WM tasks, correct recall of memory items depends on remembering them in their correct serial position, or in relation to the location where they were presented.

The centrality of binding in WM is supported by theoretical models of WM and by empirical work showing that (latent) WM capacity is strongly related to the ability to maintain bindings (Oberauer et al., 2000; Oberauer, 2005, 2009; Oberauer & Lewandowsky, 2019; Wilhelm et al., 2013). The number of bindings a person can create and maintain in WM might be the main limiting factor in WM capacity, as maintaining several bindings at the same time will increasingly lead to interference between them (Gruszka & Nęcka, 2017; Oberauer, 2009; Wilhelm et al., 2013). This upper limit on WM capacity also affects performance on WM updating tasks. That is, updating items in WM requires not just dissolving old bindings and creating new ones, but also maintaining bindings of items that should not be updated. Thus, the overlap in performance on WM updating and capacity tasks likely stems from the need in both types of tasks to create and maintain bindings in WM (Ecker et al., 2010; Frischkorn et al., 2022; Oberauer et al., 2000; Schmiedek et al., 2009; Wilhelm et al., 2013).

Nevertheless, WM updating tasks additionally require the updating of established bindings, which sets them apart from WM capacity tasks (Ecker et al., 2010; Frischkorn et al., 2022). Different updating tasks require different combinations of retrieval (making information available for immediate processing), transformation (changing a prior value into a new one, e.g., by addition or subtraction), and substitution (replacing a prior value for a new value) (Ecker et al., 2010). Ecker et al. (2010) included three measures of WM capacity as well as eight versions of a WM updating measure that required different combinations of retrieval, transformation, and substitution. After accounting for overall updating accuracy (which was positively correlated with WM capacity), they found positive correlations of around .50 between WM capacity with latent estimates of retrieval and transformation accuracy, but not with a latent estimate of substitution accuracy. Thus, when the ability to accurately substitute old with new information—a key aspect of WM updating—is sufficiently isolated from WM capacity using latent modeling, capacity and updating seem to be independent components of WM.

These findings underscore the importance of accounting for WM capacity when assessing a person's WM updating ability. This is especially important in the context of adversity research, as previous studies suggest that certain types of adverse conditions might have opposing effects on WM capacity and updating (e.g., Goodman et al., 2019; Young et al., 2018, 2022). Yet, to our knowledge, no previous research has analyzed both abilities within a single statistical model. This could lead to (1) an underestimation of the extent to which adversity undermines WM capacity, and (2) underestimation of the extent to which adversity can enhance WM updating. This, in turn, has implications for basic and applied science. For basic science, it could bias inferences about individual differences in performance on WM tasks, especially when the negative association between adversity and WM capacity is stronger than the positive association with WM updating. For applied science, it could hide from view potential pathways to leverage people's existing strengths in school or work contexts.

**Current study**

In this study, we estimated associations between latent estimates of WM capacity and updating with three types of adversity: threat, deprivation, and unpredictability. Together, these adversity types capture key dimensions in contemporary models of adversity (Ellis et al., 2009, 2022; McLaughlin et al., 2021; McLaughlin & Sheridan, 2016). Threat refers to experiences involving the potential for harm imposed by others. We focused on perceived neighborhood violence, the extent to which an individual reports having been exposed to acts of violence in their neighborhood. Deprivation refers to having a low level of resources. We focused on perceived material deprivation, a (perceived) lack of access to material resources. Unpredictability refers to variation in material deprivation over time. This definition is inspired by, but deviates from the harshness-unpredictability framework, in which unpredictability is defined as stochastic variation in harshness (age-specific rates in morbidity and mortality) over space and time (Ellis et al., 2009, 2022). We did not calculate unpredictability in neighborhood threat given that participants had at most six timepoints, and often as few as one or two, which is insufficient to calculate variation over time (Walasek et al., 2024).

We addressed three research questions. First, what is the association between adversity and WM capacity? Second, what is the association between adversity and WM updating *after* accounting for WM capacity? Third, are the directions and strengths of these associations similar or different for neighborhood threat, material deprivation, and unpredictability?

We evaluated evidence for deficit- and adaptation-based frameworks (see Figure 5.1A for a visual summary). Crucially, as deficit and adaptation processes can operate in concert (Frankenhuis, Young, et al., 2020), we could find support (or lack thereof) for both frameworks in the same model. We distinguished between three between-person data patterns: (1) lower performance, (2) higher performance, and (3) practically equivalent performance. We defined lower performance as a statistically significant negative association between a type of adversity and WM capacity or updating (irrespective of effect size). We defined higher performance as a statistically significant positive association between a type of adversity and WM capacity or updating (irrespective of effect size). We defined practically equivalent performance as an association between a type of adversity and WM capacity or updating that has a standardized effect smaller than 0.1 *and* larger than −0.1—even if the effect is statistically different from zero—which we tested using Two One-Sided T-Tests (TOST) equivalence testing (see the 'Primary analyses' section; Lakens et al., 2018).

Deficit frameworks predict a negative association between all three types of adversity and WM capacity as well as WM updating. This follows from the hypothesis that adversity leads to broad WM deficits (Farah et al., 2006; Sheridan et al., 2020). Deficit frameworks are partially supported if we find negative associations with only one (or two) types of adversity.

Within adaptation-based frameworks, theories make two predictions. First, if adaptive processes enhance WM updating and there are no impairment processes operating, we can expect a positive association between adversity and WM updating. Second, if, adaptive processes operate in concert with general impairment processes, we can expect practically equivalent WM updating performance in combination with lower WM capacity. If neither impairment nor adaptative processes are operating, we can expect both WM updating and capacity to be practically equivalent.

We also had two expectations based on prior studies. First, we expected the association between material deprivation and WM capacity to be more negative than the associations with unpredictability and neighborhood threat. This follows from findings showing that cognitive abilities are more negatively associated with cognitive deprivation than threat (Salhi et al., 2021; Sheridan et al., 2020). Although cognitive and material deprivation are distinct types of deprivation, they tend to be correlated, and are both associated with limited access to resources that stimulate cognitive development and functioning (Bradley et al., 2001; Lurie et al., 2024; Rosen et al., 2019). Therefore, we expected that their associations with WM would have comparable effect sizes. Second, researchers have hypothesized that WM updating is particularly adaptive in unpredictable and threatening environments, as it facilitates keeping track of unpredictable changes and sudden threats. Therefore, we expected WM updating to be associated with unpredictability and neighborhood threat, but not with material deprivation (Young et al., 2018; but see Young et al., 2022).

## 5.2 Methods

### Participants

Our study included 800 participant who were randomly sampled from the Longitudinal Internet studies for the Social Sciences (LISS) panel (Scherpenzeel, 2011). The LISS panel is a representative probability sample of roughly 5,000 Dutch households (~7,500 individuals) drawn from the population register by Statistics Netherlands on an invite-only basis. Households without a computer or internet connection are provided with these facilities by LISS. Each year, participants complete the same core battery of questionnaires about—among other topics—their financial situation in the past year. In addition, participants can complete additional online questionnaires every month, with variable content. The current study integrated two data sources. First, our sample of 800 participants participated in a new LISS study between October 2023 and February 2024 (hereafter referred to as 'newly collected data'), in which we included a measure of neighborhood threat and multiple measures of working memory. Second, we accessed data that were previously collected in LISS (hereafter referred to as 'the LISS archive'). See Figure 5.2 for a visual overview of the data sources and their measurement timepoints. We signed a contract with LISS stipulating that we would receive

A.

| WM capacity | WM updating | Deficit frameworks | Adaptation frameworks |
|---|---|---|---|
| Lower | Lower | Yes | No |
| Lower | Higher | Yes | Yes |
| Lower | Intact | Yes | Yes |
| Intact | Intact | No | No |

B.

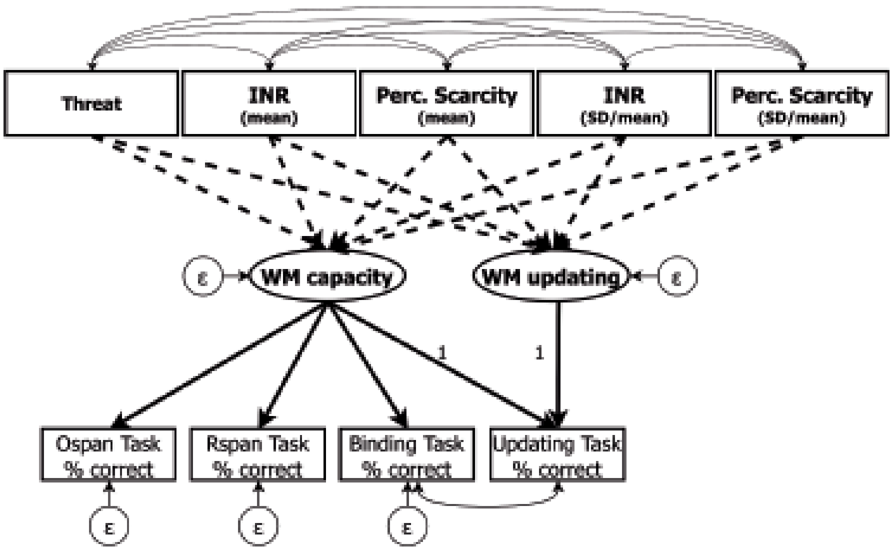



**Figure 5.1.** Overview of predictions derived from deficit and adaptation frameworks. Panel A depicts the most likely between-person data patterns based on previous literature, and whether we would consider them consistent with deficit and adaptation frameworks (see the main text for more details). Panel B depicts an overview of the preregistered Structural Equation Model. Note that this model differs slightly from the final model (see Figure 5.4). Ellipses represent latent variables, rectangles represent manifest variables, and circles represent residual variances. Unidirectional solid lines represent factor loadings, bidirectional solid lines represent covariances, and dashed lines represent regression paths. All four manifest WM measures loaded on a latent WM capacity factor, reflecting the fact that people have to hold information active in WM on all tasks. We fixed the loading of WM capacity on the Binding Task to 1, reflecting the idea that the ability to create and maintain bindings is the main limiting factor in WM capacity (Gruszka & Nęcka, 2017; Oberauer, 2009; Wilhelm et al., 2013). WM updating was modeled as a latent factor capturing the residual variance in the updating task after accounting for variance related to WM capacity. INR = income-to-needs ratio; Perc. Scarcity = perceived scarcity; SD = standard deviation.

access to the newly collected data only after Stage 1 acceptance of this Registered Report.

We based our power analysis on simulations reported by Kretzschmar & Gignac (2019), determining the required sample size to detect a small effect size ($\beta = 0.1$) with

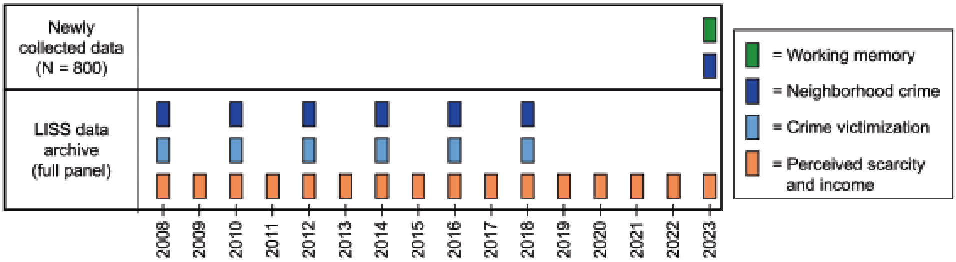



**Figure 5.2.** Overview of the different data sources used in this study. We distinguished between measures taken from the LISS data archive and measures that were newly collected in our own study between October 2023 and February 2024. Perceived scarcity and income were collected yearly in the full panel from 2008 – 2023. Neighborhood crime and crime victimization were collected across six waves between 2008 and 2018. In the newly collected data, we collected data on a measure of neighborhood threat and multiple measures of working memory. Note that participants did not have data across all timepoints of the archived studies because they joined the LISS panel more recently or because they did not participate in each wave.

at least 90% power at $\alpha = 0.05$. Assuming a reliability of at least 0.7 (which is typical for WM tasks with a number of trials similar to ours; e.g., Wilhelm et al., 2013), we required a sample size of $N = 730$. Anticipating some exclusions, we decided to include 800 participants. Participants were eligible for inclusion if they 1) were currently between 18 and 55 years old, 2) had completed at least one wave of an archived longitudinal LISS study containing measures that we use to operationalize crime neighborhood threat (see below), and 3) had given permission to link their LISS data to government microdata (not relevant here).

To ensure sufficient representation of people from lower socioeconomic backgrounds, *half* the total sample was sampled from participants who reported one or more of the following at least once in the three years: (1) a monthly income < €1,500, (2) HAVO or VWO as highest completed education (which are the two highest levels in Dutch secondary education), or (3) a score of 4 or lower on the ‘ladder of life’ (“If you imagine a ‘ladder of life’, where the first step represents the worst possible life, and the tenth (top) step the best possible life, on what step would you place yourself?”). Participants were excluded if they (1) switched to and interacted with other browser tabs *during* one or more of the cognitive tasks, (2) did not perform above chance level on the secondary processing tasks. The final sample consisted of 759 participants (Table 5.1).

**Table 5.1.** Descriptive statistics.

| Category | Statistic |
|---|---|
| Mean age (SD) | 41 (9.9) |
| Sex (% Female) | 54.4 |
| Highest completed education (%) | |
| primary school | 0.5 |
| vmbo (intermediate secondary education) | 8.3 |
| havo/vwo (higher secondary education) | 9.2 |
| mbo (intermediate vocational education) | 26.4 |
| hbo (higher vocational education) | 31.5 |
| wo (university) | 22.4 |
| other | 0.5 |
| missing | 1.2 |
| Mean number of waves (SD) | |
| INR | 13.4 (3.9) |
| Perceived scarcity | 11.1 (3.7) |
| Threat | 3.5 (1.9) |

### Measures

All independent variables, except for the income-to-needs ratio (INR) consisted of multiple items and/or scales. If all correlations between the items/scales were equal to or larger than .60 (i.e., indicating a “strong” correlation), then we computed a uniformly weighted average. If the correlation was lower than .60, we applied Principal Component Analysis (PCA) to the averaged measures and extracted only the first principal component score. We present bivariate correlations in Table 5.2, and histograms for all independent measures in the supplemental materials.

### Neighborhood threat

**Perceived neighborhood crime.** We included four items from the LISS archive collected across six waves (https://doi.org/10.17026/dans-zch-j8xt), in which participants answered how often it happens that they 1) “avoid certain areas in your place of residence because you perceive them as unsafe”, 2) “do not respond to a call at the door because you feel that it is unsafe”, 3) “leave valuable items at home to avoid theft or robbery in the street?”, 4) “make a detour, by car or on foot, to avoid unsafe areas?” on a scale of 1 (“(Almost) never”), 2 (“Sometimes”), or 3 (“Often”). We recoded these items so that 0 indicated “(Almost) never”. We then summed the responses within each wave for which participants had data, and calculated an average across the waves.

In addition, we implemented the Neighborhood Violence Scale (Frankenhuis, Young, et al., 2020; Frankenhuis & Bijlstra, 2018) in the newly collected data. The Neighborhood Violence Scale includes seven items measuring perceived exposure to neighborhood violence (e.g., "Crime is common in the neighborhood where I live"; "Where I live, it is important to be able to defend yourself against physical harm"). Participants answered these questions on a scale of 1 ("Completely disagree") to 7 ("Completely agree"). We computed an average of the seven items.

**Crime victimization.** We used data from the LISS archive collected across six waves (same dataset as above), in which participants indicated whether they fell victim to eight types of crime over the two years prior to a particular wave (0 = no, 1 = yes). We included seven items concerning exposure to crime: (1) burglary or attempted burglary; (2) theft from their car; (3) theft of their wallet or purse, handbag, or other personal possession; (4) wreckage of their car or other private property; (5) intimidation by any other means; (6) maltreatment of such serious nature that it required medical attention; (7) maltreatment that did not require medical attention. We computed a variety score by summing the exposures to *unique* types of crime across all waves. Thus, if a participants reported exposure to the same type of crime on separate waves, this counted as one exposure in the total score (Sweeten, 2012).

**Neighborhood threat composite.** We first computed an average across time for each measure separately (i.e., the two measures of neighborhood crime and the measure of crime victimization). Because correlations were below .60 (see Table 5.2), we then used PCA to extract only the first principal component score ($R^2$ = .20). The threat component was most strongly determined by the Neighborhood Violence Scale (0.63), followed by the perceived neighborhood crime scale from the LISS archive (0.40) and crime victimization (0.18).

### Material deprivation

We measured material deprivation with two separate indicators: perceived scarcity and the income-to-needs ratio.

**Perceived scarcity (mean).** We used a few items from the LISS archive that were collected on a yearly basis between 2008 and 2023 (https://doi.org/10.57990/1gr4-bf42) to index perceived scarcity. First, participants indicated how hard or easy it currently is to live off the income of their household, on a scale from 0 (very hard) to 10 (very easy). Second, participants were asked to choose which of the following best applied to their current situation: (1) "we are accumulating debt"; (2) "we are somewhat eating into savings"; (3) "we are just managing to make ends meet"; (4) "we have a little bit of money to spare"; (5) "we have a lot of money to spare". Responses were reverse-coded, so that higher scores indicated a worse financial situation. Third, participants answered which of the following issues they were confronted with at present (0 = no, 1 = yes): (1) "having trouble making ends meet"; (2) unable to quickly replace things that break"; (3) "having to lend money for necessary expenditures"; (4) "running behind in paying rent/mortgage or general utilities"; (5) "debt collector/bailiff at the door in the last month"; (6) "received financial support from family or friends in the last month".

We first computed the average across time for each item separately. Because correlations were all above .60, we calculated a uniformly weighted average.

**Income-to-needs (mean).** We calculated an income-to-needs ratio for each year using monthly self-reported net household income from the LISS archive (https://doi.org/10.57990/qn3k-as78). Zero values in household income were set to missing, as these could either indicate the lack of an income or an unwillingness to disclose the income. If monthly household income is missing (or zero) for an entire year for a participant, we used, if available, the yearly net household income they reported in the LISS archive (https://doi.org/10.57990/1gr4-bf42), dividing it by 12 to obtain a monthly estimate. First, we divided the average income per year by the *poverty threshold*, as determined by Statistics Netherlands (Van den Brakel et al., 2023; CBS, personal communication, December 15, 2023). As thresholds are only provided for households with up to three children, we applied the threshold of a household with three children to households with more than three children. Likewise, we applied the threshold of a household with two adults for households that contained three or more adults. Second, we calculated the average within-person income-to-needs ratio for each year by averaging across the monthly income-to-needs estimates.

**Unpredictability**

**Perceived scarcity (SD/mean).** This measure was based on the same items as outlined above (see Perceived scarcity (mean)). We computed unpredictability over time in perceived scarcity using the coefficient of variation, which is the within-person standard deviation across years divided by the mean (Key et al., 2017; Liu et al., 2022; Ugarte & Hastings, 2023; Walasek et al., 2024; Young et al., 2020). The mean and standard deviation in income have been found to be strongly positively correlated, indicating that people with lower incomes tend to experience less variability in income (Li et al., 2018; Young et al., 2024). For that reason, the standard deviation alone has been called into question as a measure of adversity, as the same fluctuation in income can have a greater relative impact for people close to the poverty line than for people with high incomes.

We first computed the coefficient of variation across time for each item separately. because correlations were below .60 (see Table 5.2), we then used PCA to extract only the first principal component score ($R^2$ = .38). The perceived unpredictability component was almost fully determined by the item about people's current situation (1.00), followed by difficulties to live off income (0.34) and financial troubles (0.20).

**Income-to-needs (SD/mean).** Similar to perceived scarcity, we computed unpredictability over time in the income-to-needs ratio using the coefficient of variation.

**Table 5.2.** Spearman correlation between the main independent variables.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. INR (M) | - | | | | | | | | | | | | | |
| 2. Living off income (M) | -0.52*** | - | | | | | | | | | | | | |
| 3. Financial troubles (M) | -0.43*** | 0.71*** | - | | | | | | | | | | | |
| 4. Current situation (M) | -0.51*** | 0.75*** | 0.69*** | - | | | | | | | | | | |
| 5. Perceived scarcity (M) | -0.55*** | 0.96*** | 0.79*** | 0.89*** | - | | | | | | | | | |
| 6. INR (CV) | -0.17*** | 0.15*** | 0.21*** | 0.14*** | 0.18*** | - | | | | | | | | |
| 7. Living off income (CV) | 0.19*** | -0.20*** | -0.02 | -0.18*** | -0.19*** | 0.15*** | - | | | | | | | |
| 8. Financial troubles (CV) | -0.36*** | 0.64*** | 0.92*** | 0.60*** | 0.71*** | 0.24*** | 0.05 | - | | | | | | |
| 9. Current situation (CV) | 0.20*** | -0.11** | 0.04 | -0.11** | -0.11** | 0.12** | 0.34*** | 0.13*** | - | | | | | |
| 10. Perceived scarcity (CV) | 0.21*** | -0.11** | 0.05 | -0.16*** | -0.11** | 0.18*** | 0.35*** | 0.16*** | 1.00*** | - | | | | |
| 11. Neighborhood safety | -0.13*** | 0.19*** | 0.16*** | 0.17*** | 0.20*** | 0.05 | -0.10* | 0.12** | -0.05 | -0.05 | - | | | |
| 12. Neighborhood Violence Scale | -0.22*** | 0.26*** | 0.19*** | 0.22*** | 0.27*** | 0.02 | -0.10* | 0.16*** | -0.06 | -0.05 | 0.24*** | - | | |
| 13. Crime victimization | 0.01 | 0.12** | 0.18*** | 0.16*** | 0.15*** | 0.10** | 0.01 | 0.17*** | 0.07 | 0.07 | 0.06 | 0.12** | - | |
| 14. Threat | -0.21*** | 0.29*** | 0.24*** | 0.26*** | 0.31*** | 0.07 | -0.12** | 0.20*** | -0.06 | -0.04 | 0.58*** | 0.89*** | 0.26*** | - |
| Mean | 1.99 | 4.17 | 1.30 | 2.35 | 2.60 | 0.22 | 0.27 | 0.21 | 0.27 | -0.01 | 1.45 | 2.39 | 1.04 | -0.02 |
| SD | 0.76 | 1.60 | 0.53 | 0.75 | 0.87 | 0.19 | 0.17 | 0.24 | 0.15 | 0.99 | 1.47 | 0.95 | 1.27 | 0.68 |
| Min | 0.09 | 1.00 | 1.00 | 1.00 | 1.00 | 0.01 | 0.00 | 0.00 | 0.00 | -1.86 | 0.00 | 1.00 | 0.00 | -1.07 |
| Max | 6.10 | 11.00 | 4.44 | 5.00 | 5.93 | 1.52 | 0.99 | 0.92 | 0.93 | 4.42 | 8.00 | 6.86 | 7.00 | 3.68 |
| Skew | 1.06 | 0.76 | 2.47 | 0.44 | 0.87 | 2.31 | 0.95 | 0.62 | 0.22 | 0.34 | 1.18 | 1.33 | 1.28 | 1.21 |
| Kurtosis | 3.42 | 1.39 | 6.86 | -0.08 | 1.08 | 8.83 | 1.22 | -1.01 | 0.87 | 1.34 | 1.22 | 2.35 | 1.27 | 2.05 |

*Note:* * = $p$ < .05, ** = $p$ < .01, *** = $p$ < .001. CV = coefficient of variation, INR = income-to-needs ratio, M = mean, Perc. Scarcity = perceived scarcity

**WM tasks**

The WM tasks were all part of the newly collected data. All materials and scripts for the cognitive tasks can be found at https://stefanvermeent.github.io/liss_wm_profiles_2023/materials/README.html. Prior to collecting LISS data, we conducted a pilot study among in a Dutch sample (*N* = 100) through Prolific Academic. The main goals of this pilot study were to collect participant feedback (e.g., difficulty of instructions, whether we included sufficient breaks) and to analyze performance and correlations between tasks. The results of this pilot study are described in more detail in the Supplemental Materials (https://stefanvermeent.github.io/liss_wm_profiles_2023/supplement/README.html).

**Operation Span Task.** The Operation Span Task (Figure 5.3A) is a common measure of WM capacity (Conway et al., 2005; Wilhelm et al., 2013). In this task, participants alternate between a primary memorization task and a secondary processing task. On each trial, the task is to memorize a sequence of letters in the correct order (from a set of 12 letters). Each letter is presented for 1,000 ms in the center of the screen. Next, participants see a simple mathematical equation including the outcome. Their task is to indicate whether the outcome is correct or incorrect by pressing either the 'a' or 'l' key on their keyboard. The equations always contain one addition or subtraction, with numbers ranging between one and 10. Outcomes are always positive integers. On each trial, participants have to memorize between four and six letters, with each set size repeated three times. At the end of each sequence, all letters are presented in a 3×4 grid, and participants click the letters in the correct order.

Participants first practiced the letter task (three times), then the math task (eight times), and then the full task (three times). If they performed at or below chance, they had the opportunity to either repeat a part or advance to the next part. After practicing, participants completed 9 test trials, with a total of 45 recall items and 45 math items. We computed an operation span score by calculating the proportion of letters recalled in the correct sequential position across trials (Conway et al., 2005).

**Rotation Span Task.** The Rotation Span Task (Figure 5.3B) is similar to the Operation Span Task and was adopted from Wilhelm et al. (2013). On each trial, the task is to memorize the orientation of a sequence of arrows in the correct order. Arrows can take on eight different orientations, with increments of 45∘. Each arrow is presented for 1,000 ms in the center of the screen. Next, participants see a capital 'G' or 'F' that is rotated at one of eight different orientations, with increments of 45∘. Their task is to indicate whether the letter is mirrored or not. On each trial, participants have to memorize between two to five arrows, with each set size repeated three times. At the end of each sequence, all arrows are presented simultaneously, and participants click the arrows in the correct order.

Participants first practiced the arrow task (three times), then the letter task (eight times), and then the full task (three times). If they performed at or below chance, they had the opportunity to either repeat a part or to advance to the next part. After practicing, participants completed 12 test trials, with a total of 45 recall items and 45 letter items. We computed a rotation span score by calculating the proportion of arrows recalled in the correct sequential position across trials (Conway et al., 2005).

**Binding-Updating Task.** The Binding-Updating task (Figure 5.3C) was adopted from Wilhelm et al. (2013). On each trial, participants see a 3×3 grid, with a fixation cross in the central cell. After 1,000 ms, they are presented with a sequence of numbers (0-9) in random locations of the grid. Each new number is presented for 1,500 ms, after which it disappears for 500 ms before the next number is presented. The task is to remember the last number they see in each location. Memory set sizes (i.e., the number of unique locations in the grid) ranges between three and five. On half of the trials, only one number is presented in each location. These constitute the binding trials. On the other half of the trials, some letters are presented in the same location as previous numbers, requiring mentally replacing the old number with the new number. These constitute the updating trials. We use two, three, and four updating steps, each repeated in combination with the different set sizes. At the end of the trials, participants indicate which letter they saw last in each location in random order.

Participants first completed four practice trials. If they performed at or below chance, they had the opportunity to either repeat the practice trials or to advance to the actual task. After practicing, they completed 18 test trials, of which nine were binding-only (24 recall items in total) and nine were updating trials (24 recall items in total). We computed a binding score by calculating the overall recall accuracy (%) across trials with zero updating steps. We computed an updating score by calculating the overall recall accuracy (%) across trials with one or more updating steps.

### Procedure

We received ethical approval from the Ethics Review Board of the Faculty of Social & Behavioral Sciences of Utrecht University (FETC20-490). Upon starting the study, participants were informed that the study could only be completed on a laptop or desktop PC. If they attempted to start the study on a tablet or smartphone, they were unable to advance and prompted to switch to a suitable device. Participants started with the WM tasks, which on average took between 20 and 25 minutes. The WM tasks were completed in fullscreen mode. If participants left fullscreen mode at any moment during the tasks, they saw instructions at the top of their screen that allowed them to return to fullscreen mode. The order of the WM tasks was counterbalanced, and participants had the opportunity to take breaks at regular intervals.

After the cognitive tasks, participants answered three questions about the environment in which they completed the WM tasks: 1) “How much noise was there in your

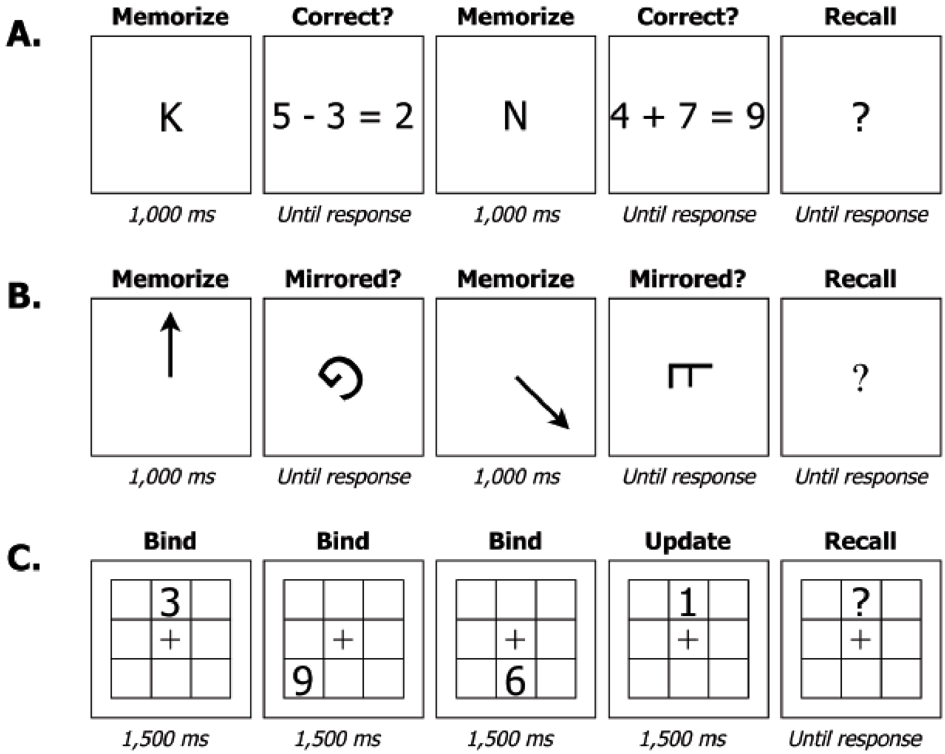



**Figure 5.3.** Overview of the working memory tasks. Panel A: Operation Span Task. Participants memorized letters in the correct order, while engaging in a secondary math task. Panel B: Rotation Span Task. Participants memorized the orientation of arrows, while judging whether letters were mirrored or normal in a secondary task. Panel C: Participants memorized numbers in the correct location in a 3×3 grid. On half of the trials, all numbers were presented in unique locations, only requiring binding the numbers to the correct position. On the other half, some numbers were presented in the same location as a previously presented number, requiring updating. Note: stimuli are not to scale.

environment during the memory tasks?"; 2) "Were you at any moment interrupted during the memory tasks?"; 3) "Did you at any moment during the memory tasks leave the computer?". Next, they completed questionnaires about their future orientation (not considered here), personality (not considered here), past adversity exposure, and recent adversity exposure. Finally, they completed a standard set of evaluation questions asking about their experiences with the study, with the possibility to provide open-ended feedback. This part on average took 5 minutes. Participants received €7.50 for their participation through LISS. If participants experienced difficulties of any sort, they could contact the LISS helpdesk.

### Proposed analysis plan

The Stage 1 protocol of this Registered Report can be found at https://osf.io/dp7wc.

**Data access**

The working memory data and one of the neighborhood threat indices were collected between October 2023 and February 2024, prior to submitting the Stage 1 protocol. These data were only made available to the first author after Stage 1 acceptance, as stipulated in a signed contract with LISS. During planning of the study, the first author accessed the LISS data archive and inspected three waves of the LISS data containing the items about neighborhood safety and crime exposure, as well as the three most recent monthly data collections containing basic demographic info. The purpose was to ascertain the number of individuals who had finished the previous waves in the LISS data archive and were presently still participating in the panel (i.e., to see if we could reasonably create a link between the LISS data archive and newly collected data).

All data access events were automatically detected and logged on the GitHub repository using the *projectlog* R package (Vermeent, 2023). We took the following measures to prevent bias: 1) we randomly shuffled the participant IDs in each data set using the *projectlog* R package, so that we were unable to link participant data between (waves of) studies in the LISS data archive; 2) we did not inspect any of the measures that will be part of our adversity composites; 3) we did not know which participants would be selected for the newly collected data; 4) the primary analyses will be based on composite measures that combine measures from the LISS data archive with measures from the newly collected data.

**Primary analyses**

See Figure 5.1B for an overview of the model specification. We fitted a single model containing all adversity measures using the *lavaan* R package (Rosseel, 2012). We used robust maximum likelihood estimation to account for non-normality. Missing data were handled using full information maximum likelihood (FIML). We accounted for clustering within families using the *lavaan.survey* R package (Oberski, 2014).

WM capacity was estimated as a latent factor loading on all outcome measures. In addition, we estimated WM updating as a latent factor capturing residual variance in the updating measure. Thus, this factor accounted for updating-specific variance after accounting for WM capacity. We estimated the effect of each adversity type (dashed lines in Figure 5.1B) through regression analyses. Each association was controlled for: (1) age in years ; (2) the quadratic effect of age; (2) environmental noise ("How noisy was your environment during the memory tasks", rated on a scale of 1 (very little noise) to 5 (a lot of noise)); (3) two items measuring interruptions ("Where you at any moment interrupted during the memory tasks?" and "Did you at any moment during the memory tasks leave your computer?", rated as yes or no). Goodness of fit was assessed using the comparative fit index (CFI) and the root mean square error of approximation (RMSEA). CFI values > .90 and RMSEA values < .08 were interpreted as acceptable model fit, and CFI values > .95 and RMSEA values ≤ .06 as good model fit (Hu & Bentler, 1999).

We anticipated that we may have to optimize the model further in case of bad model fit, and therefore planned to estimated the model in two steps to prevent bias. First, we constructed the measurement model of WM, without including the adversity measures. This step was planned to be carried out prior to accessing any of the adversity measures. Once we obtained at least acceptable model fit, we accessed and added the adversity measures to the model. This procedure was tracked and timestamped on the GitHub repository using the procedure outlined above. We controlled for multiple testing across adversity measures, but separately for each outcome measure, using the false discovery rate (Benjamini & Hochberg, 1995; Cribbie, 2007).

To statistically test whether small effects were practically equivalent to zero we used Two One-Sided T-tests (TOST) equivalence testing (Lakens et al., 2018), using −0.1 and 0.1 as equivalence bounds. TOST equivalence testing allows us to conclude practically equivalent performance based on a significant effect, rather than erroneously interpreting a non-significant effect as evidence for the absence of an effect. We considered any effect that fell within this region to reflect practical equivalence, that is, a between-person difference in performance that is practically equivalent to zero. TOST provides two *p*-values, one testing against the upper bound and one testing against the lower bound; we report only the largest of the two *p*-values.

## 5.3 Results

### Confirmatory analyses

#### Model fit

The preregistered measurement model specification did not converge. A model version excluding the covariance between manifest binding and updating did converge, but resulted in suboptimal fit (Robust CFI = 0.95, robust RMSEA = 0.12, 95% CI = [0.09, 0.14]). Modification indices indicated that model fit would improve most from estimating the covariance between Rotation Span and Operation Span, which is in line with previous factor models of working memory containing span tasks as a subset of other working memory tasks (e.g., Löffler et al., 2024). A model incorporating an estimate of this covariance provided a good fit to the data (Robust CFI = 1, robust RMSEA = 0, 95% CI = [0, 0]). After finalizing the measurement model, we constructed the final structural model by adding all predictors and covariates to the model, which resulted in a good model fit (Robust CFI = 0.99, robust RMSEA = 0.03, 95% CI = [0, 0.03]). Figure 5.4 presents a visual overview of the final model.

#### Associations between adversity and WM

The main results of the associations between the adversity measures and WM are summarized in Figure 5.5. None of the adversity measures were significantly associated with WM capacity after adjusting for multiple testing (all *ps* ≥ .063). We also did not find evidence for practical equivalence for associations between any of the adversity

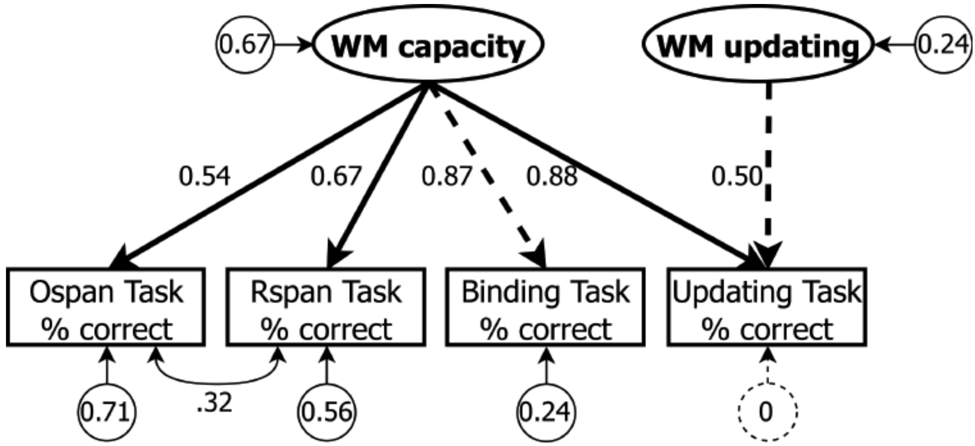



**Figure 5.4.** Overview of the final measurement model of WM performance. Ellipses represent latent variables, rectangles represent manifest variables, and circles represent unstandardized residual variances. Unidirectional lines represent standardized factor loadings and bidirectional lines represent covariances. All four manifest WM measures loaded on a latent WM capacity factor, reflecting the fact that people have to hold information active in WM on all tasks. We fixed the loading of WM capacity on the Binding Task to 1, reflecting the idea that the ability to create and maintain bindings is the main limiting factor in WM capacity (Gruszka & Nęcka, 2017; Oberauer, 2009; Wilhelm et al., 2013). WM updating was modeled as a latent factor capturing the residual variance in the updating task after accounting for variance related to WM capacity. WM = working memory; Ospan = Operation Span; Rspan = Rotation Span.

measures and WM capacity (all *ps* ≥ .055). Similarly, none of the adversity measures were significantly associated with WM updating after adjusting for multiple testing (all *ps* ≥ .370). We also did not find evidence for practical equivalence to zero for associations between any of the adversity measures and WM updating (all *ps* ≥ .109).

**Posthoc non-preregistered analyses**

We conducted three posthoc non-preregistered analyses, described in more detail in the supplemental materials. First, to contextualize our findings based on latent WM estimates, we estimated associations between adversity and performance on the separate WM tasks using four linear regressions. Threat had small, significant negative associations with performance on the Rotation Span Task ($\beta = -0.13$, $p = .002$), Operation Span Task ($\beta = -0.14$, $p = .002$), and Binding Task ($\beta = -0.12$, $p = .004$). None of the types of adversity were significantly associated with performance on the Updating Task (all *ps* > .181), and only the association with unpredictability in the income-to-needs was practically equivalent to zero ($p = .041$).

Second, the inconclusive nature of our confirmatory results could indicate that the true effect sizes were smaller than the effect size of interest that we used for our power analysis ($\beta = 0.1$; i.e, that we lacked sufficient power). To explore this, we con-

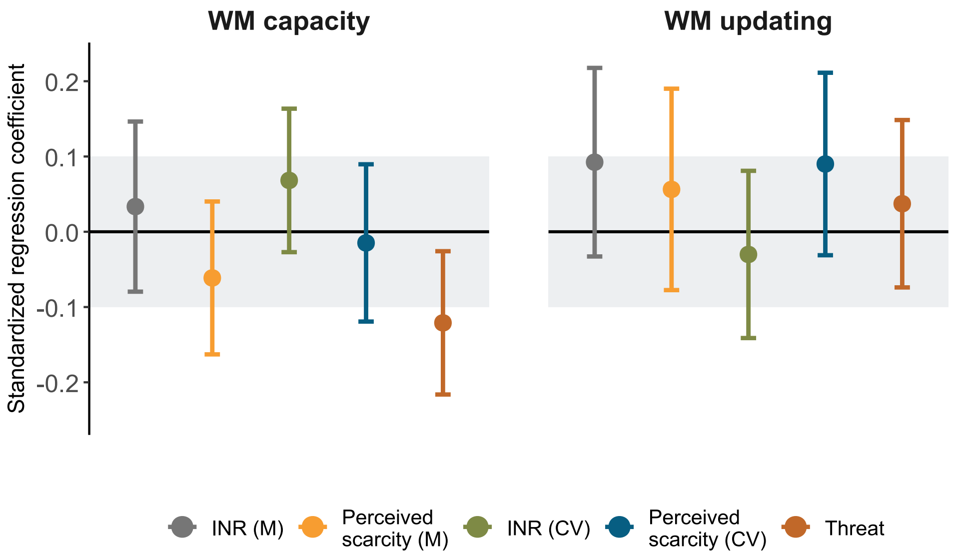



**Figure 5.5.** Results of the structural part of the SEM model testing the association between threat, deprivation, and unpredictability on latent estimates of WM capacity and WM updating. The gray area shows the area of practical equivalence. Solid points indicate effects outside the area of practical equivalence, which was true for all effects. Standard errors represent the 95% confidence intervals. CV = coefficient of variation; INR = income-to-needs ratio; M = mean; WM = working memory.

ducted an alternative test for the absence of an association between adversity and WM by constraining regression paths between adversity and WM factors to zero in the SEM. Constraining all paths to latent WM capacity to zero significantly reduced model fit, although the change in AIC was below the cut-off as proposed by Burnham & Anderson (2002), $\Delta$ AIC = 7.62, $\Delta\ \chi(5) = 14.20$, $p = .014$, Robust CFI = 0.99, robust RMSEA = 0.03, 95% CI = [0.01, 0.04]. Constraining all paths to latent WM updating did not significantly reduce model fit, $\Delta$ AIC = 3.81, $\Delta\ \chi(5) = 5.85$, $p = .321$, Robust CFI = 0.99, robust RMSEA = 0.03, 95% CI = [0, 0.03]. Thus, these results were somewhat inconsistent with the preregistered frequentist equivalent tests.

Third, as a non-preregistered robustness check, we calculated Bayes factors for the preregistered equivalence tests using the *bain* package (Hoijtink et al., 2019), in which we evaluated whether the observed data are more likely under the hypothesis that the effects fall within the equivalence bounds, relative to the hypothesis that the effects fall outside of the equivalence bounds. The results are summarized in Table A4.3. For all but one association, the model comparisons showed at least strong evidence in favor of the data being more likely under the hypothesis that the effects fell within the equivalence bounds ($BF_{10}$ ranging between 16.9 and 158.9. The only exception was the association between threat and WM capacity, for which we found moderate evidence in favor of the data being more likely under the hypothesis that the effect fell within the equivalence bounds ($BF_{10}$ = 5.5. Thus, these results were inconsistent with the preregistered frequentist equivalent tests, which did not find evidence for practical equivalence.

### Deviation from the Stage 1 protocol

In the Stage 1 protocol, we planned to first access the dependent variables to construct the SEM, and then access the independent variables. Due to an unintended error, the first author already accessed the datasets containing the measures that would be used to compute the independent variables before finalizing the SEM. However, beyond reading them into the R environment, these data were not yet inspected, manipulated, or summarized. We contacted the PCI recommender upon finding out about this deviation, and agreed to describe this deviation as done here. For the sake of transparency, we timestamped the scripts for processing the independent variables at the moment of this unintended data access (https://github.com/StefanVermeent/liss_wm_profiles_2023/blob/d143e55/scripts/2_pipeline/1_ivs.R). They contain the code to read in the data, but no code yet for any type of data cleaning or variable computation.

## 5.4 Discussion

We investigated associations between adversity (threat, material deprivation, and unpredictability) and WM capacity, a person's ability to hold information available for later processing, as well as WM updating, a person's ability to mentally replace old with new information. We distinguished between WM capacity and updating on a latent level using four different tasks, three of which are primarily construed as WM capacity tasks, and one that is primarily construed as a WM updating task. The WM capacity factor loaded on performance of all four tasks, in line with previous findings (Frischkorn et al., 2022; Gruszka & Nęcka, 2017; Oberauer, 2009; Wilhelm et al., 2013). An additional WM updating factor accounted for the portion of variance in the Updating Task that was not explained by WM capacity. We did not find any consistent associations between adversity and WM capacity nor updating in our preregistered analyses. On the one hand, none of the associations significantly differed from zero. On the other hand, none of the associations fell within the pre-specified region of practical equivalence to zero (i.e., a between-person difference in performance that is practically equivalent to zero).

The confirmatory results were not consistent with hypotheses generated from a deficit framework. A large literature has documented negative associations between exposure to early-life adversity—especially deprivation—and WM capacity, which persists into adulthood (Farah et al., 2006; Goodman et al., 2019; Sheridan et al., 2022; Sheridan & McLaughlin, 2014; Young et al., 2018; but see Nweze et al., 2021). Similarly, studies with young adults have found that a higher frequency of recent as well as lifetime stressful major life events (i.e., negative events with a clear onset and offset, unlike chronic adversity) is also negatively associated with WM capacity (Klein & Boals, 2001;

Shields et al., 2017, 2019). The results were also not consistent with hypotheses generated from adaptation frameworks. Recently, a small set of studies documented intact and even higher WM updating performance in adolescents and adults who reported more exposure to childhood adversity (Young et al., 2018, 2022). These associations have been interpreted as reflecting developmental adaptations to adversity: in more threatening and unpredictable environments, it may be beneficial to be able to rapidly update the items held in WM (Ellis et al., 2017, 2022; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013). In contrast, we did not find consistent associations between adversity exposure and WM updating. These findings are inconclusive, as we also did not find evidence for practical equivalence in our preregistered analysis.

A set of non-preregistered robustness checks were comparatively more consistent with practically equivalent performance, although they did not fully rule out the existence of small associations between adversity exposure and working memory performance. First, A Bayesian reanalysis of the preregistered equivalence tests (using the same equivalence bounds) provided strong evidence in favor of the hypothesis that working memory performance was practically equivalent, in contrast to the preregistered analyses. Second, constraining the regression paths in the SEM to zero somewhat reduced model fit for WM capacity, but not for WM updating. This suggests that there may have been systematic associations with WM capacity that were smaller than the equivalence bounds used in the (Bayesian) equivalence tests. If true, the associations would be smaller than we expected based on the literature outlined above, and would require a larger sample size to reliably detect. These analyses were not part of the registered analysis protocol, and therefore should be interpreted with sufficient caution pending replication.

The Updating Task shared a large proportion of variance with the WM capacity measures, which aligns with prior psychometric work focused on the structure of WM (Lewandowsky & Farrell, 2010; Oberauer et al., 2000; Wilhelm et al., 2013). This highlights an important methodological issue for the field of adversity research, especially researchers working from adaptation frameworks, who hypothesize distinct effects of adversity on different components of WM (in contrast to deficit-oriented researchers, who expect adversity to have a negative effect on all components of WM). Specifically, adaptation-oriented researchers have hypothesized that certain types of adversity may enhance WM updating through developmental adaptation, while impairing WM capacity (Ellis et al., 2022; Young et al., 2018, 2022). So far, this hypothesis has—to our knowledge—only been tested based on raw performance on single WM updating tasks. However, if true, performance on single WM updating tasks may substantially underestimate positive associations between adversity and WM updating, as raw performance may be influenced by both deficit and adaptation processes (the former influencing WM capacity, inadvertently measured in WM updating tasks). Leveraging these psychometric insights will be pivotal to better understanding associations between adversity and WM for future studies.

Aside from psychometric considerations, a second potential reason for the discrepancy between our findings and those from previous studies is that our investigation focused on adverse experiences in adulthood. In contrast, most previous studies have focused on the effects of either childhood adversity or stressful life events. It is possible that, relative to childhood adversity, the association between adversity in adulthood and WM varies as a function of other factors. For example, the association between adversity in adulthood and WM might be stronger for people who also experienced adversity during childhood, either due to early developmental calibration to chronic stress and/or due to greater lifetime exposure to stress (Hostinar & Gunnar, 2013; Shields et al., 2017).

### Strengths, limitations, and future directions

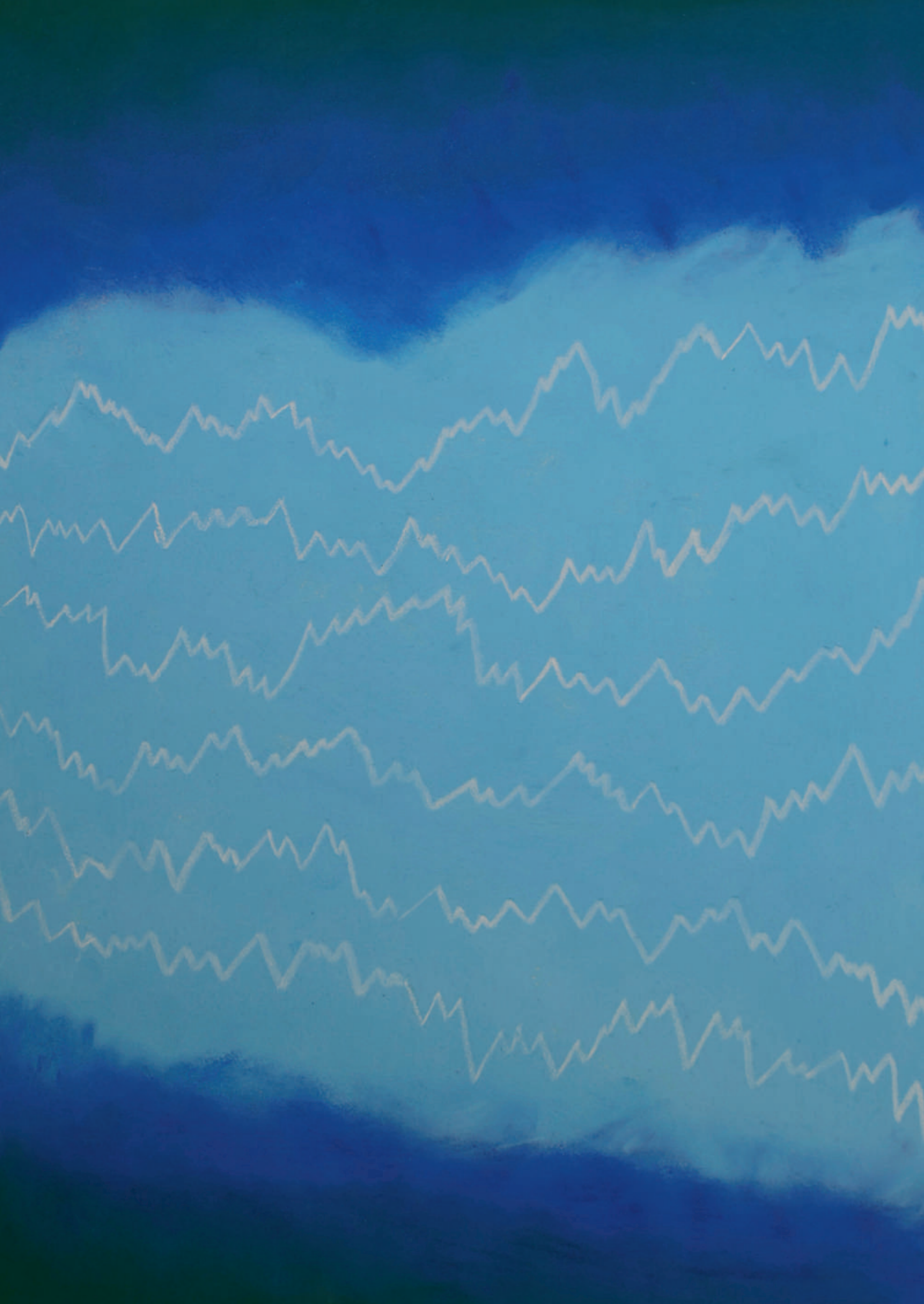
This study had several strengths. First, the sample was drawn from the Dutch LISS panel, which provides a large, representative sample of the Dutch population. Second, we drew on the longitudinal nature of the LISS panel to estimate three key dimensions of adversity exposure (threat, deprivation, and unpredictability), using several indicators for each. Third, we included four WM tasks, and used SEM to separate variance related to WM capacity from variance related to WM updating. This allowed us to more precisely estimate capacity and updating as two key components of WM.

This study also had limitations. First, WM updating was measured as the residual variance of a single task after accounting for WM capacity. This means that the latent WM updating measure was not a pure measure of WM updating, but also included measurement error. This decision was mainly guided by the limited number of tasks that could be included due to time constraints. To obtain a more reliable measure of WM updating, it would be better to include several different WM updating tasks, just like we used several different WM capacity tasks. Second, as this was an online study, we had only limited control over the environment in which people completed the study. The models accounted for self-reported noise and distractions, and we excluded participants who interacted with other browser tabs during the WM tasks. Yet, there may have been other, unmeasured factors that could lower the reliability of our study relative to lab-based studies. Third, our results appeared to be underpowered, despite including 759 participants, which suggests that the associations between adversity and WM in adulthood are smaller than expected based on previous literature. Finally, our study did not include genetic measures. It is well-established that genetic variation accounts for a substantial portion of the individual differences in executive functions (Friedman et al., 2008). However, for genetics to have confounded our study, it would need to have caused both individual differences in cognition and in adversity exposures—producing non-causal associations between adversity and cognition. Testing this fuller picture would require using genetically informative designs.

Future research could build on the current study in four ways. First, modeling WM ability on a latent level using multiple tasks could be applied more broadly in the field of adversity research, as studies rarely directly account for the overlap in key cognitive processes across WM tasks WM tasks. This is especially important for adaptation-based research focusing on WM updating ability, as WM capacity plays a substantial role in performance on updating tasks. Second, future work is needed to better understand the role of developmental timing: is adversity experienced earlier or later in life associated differently with WM across the lifespan? Third, more research is needed to better understand the relationship between more objective (e.g., income-to-needs ratio) and subjective (e.g., perceived scarcity) indicators of adversity, as well as their respective association with cognitive functioning (Smith & Pollak, 2021). In our study, mean INR and mean perceived scarcity correlated moderately, suggesting that they capture similar but separable aspects of material deprivation, which could show different associations with cognition. Fourth, the field needs to account for functional heterogeneity within adversity-exposed populations (Masten, 2001). In a recent study, the majority of U.S. adolescents with low socioeconomic resources performed on par with their privileged peers (Shariq et al., 2024). The deficit pattern observed in the population as a whole was driven by a much smaller, cognitively less resilient, subgroup. A valuable direction is to combine such a 'person-centered' approach with structural equation modeling to estimate specific WM abilities among different subgroups within adversity-exposed populations.

## Conclusion

Over the last decade, adversity research has been shifting towards a more balanced view, focusing not just on cognitive deficits but also on potential adaptations. This has spurred a growing number of studies investigating more precise links between specific types of adversity and different cognitive abilities. Adaptation perspectives in particular have emphasized the need to be more precise about how specific types of adversity are associated with specific cognitive abilities. However, this increased need for precision in the measurement of cognitive abilities requires more advanced psychometric approaches. For this, adversity researchers can draw, more than they currently do, on decades of psychometric research focused on WM and other cognitive abilities. Here, our psychometric investigation of WM yielded inconclusive associations with adverse experiences in adulthood. Building on this work will ultimately lead to a better understanding of the unique abilities that develop in contexts of adversity, as well as more precise intervention targets.



# Chapter 6

## General discussion

## 6.1 Fitting the pieces together

In the preceding chapters, I apply a methodological approach—grounded in Drift Diffusion Modeling (DDM) and structural equation modeling—to measure executive functioning (EF) abilities in people exposed to adversity. Using DDM, I translate raw performance into three distinct cognitive processes: the speed of evidence accumulation (drift rate), response caution (boundary separation), and speed of stimulus encoding and response execution (non-decision time). Using structural equation modeling, I investigated the extent to which cognitive processes are task-general (shared across tasks) or task-specific (unique to particular tasks). I investigate associations between these cognitive processes and exposure to three types of adversity: threat (Chapters 2-5), material deprivation (Chapters 2, 3, and 5), and unpredictability (Chapters 4 and 5). In addition, I investigate associations across different developmental stages, focusing on middle childhood (Chapter 2), young adulthood (Chapter 4), and adulthood (Chapters 3 and 5).

Taken together, I show that adversity researchers analyzing raw performance (e.g., mean response time, accuracy) will overestimate the association between adversity exposure and specific EF abilities. This general conclusion is based on three key findings. The first key finding, supported by Chapters 2-4 (but not Chapter 5), is that people with more exposure to adversity show lower task-general processing speed (as measured using DDM's drift rate parameter). That is, they respond more slowly largely due to cognitive processes that are shared across different EF tasks. The second key finding, also supported by Chapters 2-4, is that after accounting for task-general processing speed, the specific EF abilities of people with more exposure to adversity do not appear to be lower (or higher) than those of people with less exposure to adversity. In fact, in Chapter 2, five out of six associations with task-specific drift rates were practically equivalent to zero, suggesting intact processing. The third key finding, supported by Chapters 2 and 4 (but not Chapter 3), is that people with more exposure to adversity use different strategies on EF tasks. Specifically, I find that children with more exposure to household threat (but not material deprivation) respond more cautiously (as measured with DDM's boundary separation parameter), and that young adults with more exposure to childhood threat and unpredictability process information more holistically.

## 6.2 Key finding 1: Adversity exposure is associated with lower task-general processing speed

In the preceding chapters, I interpret the negative association between adversity exposure and task-general drift rate as reflecting lower general speed of processing. This interpretation follows from specific patterns observed in these studies, and aligns with previous literature. In Chapter 2, the task-general drift rate loaded equally strongly on drift rates of EF tasks as well as a basic processing speed task. Similarly, in Chapter 3, loadings were comparable across experimental conditions of EF tasks (i.e., the switch or incongruent condition, requiring EF ability), non-experimental conditions (i.e., the repeat or congruent condition, not requiring EF ability), and a basic processing speed task (not requiring EF ability). In Chapter 4, performance differences on the Flanker task were explained mostly by the strength of perceptual processing, not by the ability to inhibit distractions. Combined, these results support the view that task-general drift rate captures processes that are not unique to specific EF tasks.

The processing speed interpretation of task-general drift rate is consistent with recent studies applying DDM to EF tasks (Frischkorn et al., 2019; Hedge et al., 2022; Löffler et al., 2024). These studies found that speed of processing mostly—and in some cases fully—explained shared variance in drift rates across EF tasks. One study found that task-general drift rate correlated only moderately with a general intelligence factor ($r$ = .43) and a working memory capacity factor ($r$ = .41), while the latter two correlated more strongly with each other ($r$ = .76) (Löffler et al., 2024). This suggests that the task-general drift rate factor of EF tasks is related to, but conceptually distinct from, general intelligence and working memory capacity.

However, a competing interpretation is that shared variance among EF tasks represents executive attention, which refers to a general ability to focus on task-relevant information while ignoring irrelevant distractions (Mashburn et al., 2023; Zelazo & Carlson, 2023). Specifically, executive attention is thought to support a person's ability to *maintain* information in working memory for immediate processing, and to *disengage* from information that is no longer relevant (Burgoyne & Engle, 2020; Shipstead et al., 2016). Executive attention can offer a mechanistic explanation for the general factor that accounts for variance across many cognitive tasks, often referred to as general intelligence or $g$ (Burgoyne et al., 2022). Thus, shared variance across EF tasks could reflect general executive processes, rather than basic processing speed. A similar argument is made by *process overlap theory*, which states that the general factor reflects a shared dependence on general executive processes (Kovacs & Conway, 2016). Importantly, process overlap theory does not consider the general factor to be a unitary cognitive process that *causes* differences in specific abilities. Rather, the general factor arises as a statistical artifact as specific cognitive abilities draw from a shared set of general processes (Kovacs & Conway, 2019). This is an important distinction: while the task-general factor may resemble a unitary process in latent models, it could actually arise from a combination of (partially) independent processes.

Task-general drift rate could similarly reflect several processes. Differences in drift rates might reflect a combination of task-specific processes (e.g., EF abilities), state factors (e.g., motivation, fatigue), and trait factors (e.g., general speed of processing, functional or structural brain differences) (Weigard & Sripada, 2021). At the same time, task-general drift rate appears to be more stable over time (Schubert et al., 2016;

Weigard et al., 2021). It also has better convergent validity, correlating, for instance, with self-report measures of self-control, which is conceptually similar to EF (Weigard et al., 2021). However, these studies did not account for adversity exposure. It is possible that task-general drift rate is more strongly influenced by state factors for people with more adversity exposure. Thus, processing speed may be but one potential explanation for the negative associations we observed between adversity exposure and task-general drift rate; other explanations may include motivation, stress, fatigue, and a strategic deployment of cognitive resources. This also means that lower task-general drift rate does not necessarily (only) reflect a cognitive deficit.

More work is needed to better understand why exposure to adversity is negatively associated with task-general drift rate. To the extent that it reflects basic processing speed, it could partially be the result of structural and/or functional brain changes, like reduced white matter tract integrity (Fuhrmann et al., 2020; Kievit et al., 2016). White matter tracts support information processing and communication between key networks involved in EF (Ribeiro et al., 2024). Childhood exposure to threat and deprivation has been associated with reduced white matter tract integrity (McLaughlin et al., 2019). Changes in white matter associated with childhood adversity appear to persist into adulthood (McCarthy-Jones et al., 2018), which could explain the associations between childhood adversity and task-general drift rate in adulthood in Chapter 3. Relatedly, early exposure to cognitive deprivation (i.e., a lack of cognitive stimulation) disrupts the development of basic sensory and perceptual processes, which can have negative downstream effects on the development of EF (Rosen et al., 2019). Yet, task-general drift rate may at least partly reflect processes that are more context-dependent, such as EF engagement or task familiarity (Niebaum & Munakata, 2023), rather than basic processing speed. Such processes may be more malleable than basic processing speed, e.g., through task manipulations that increase the familiarity of content, or that make people more willing to exert effort. This could make them valuable targets for interventions.

## 6.3 Key finding 2: Adversity exposure is not associated with specific EF abilities

A second consistent finding throughout my dissertation is that after controlling for task-general processing speed, adversity exposure is not associated with specific EF abilities, as measured with drift rates. In Chapter 2, I show that children with more exposure to household threat in the preceding year exhibit intact drift rates on an inhibition, attention shifting, and mental rotation task, after accounting for lower task-general processing speed. In addition, material deprivation is associated with intact drift rates on an inhibition and mental rotation task, as well as intact task-general processing speed. Chapter 3 paints a similar, but more nuanced picture, by including two inhibition tasks, three attention-shifting tasks, and a basic processing speed task. After accounting for task-general processing speed, adversity is negatively associated with several task-specific drift rates, particularly effects of childhood threat on attention-shifting tasks. However, the correlations between these task-specific drift rates are low, even between tasks that are thought to measure the same EF ability. Thus, it appears that they do not capture inhibition or attention-shifting ability, but rather more unique features of individual tasks. In Chapter 4, which focuses on the Flanker task, young adults' exposure to childhood threat and unpredictability is not associated with inhibition ability. Rather, their lower performance on the task is mostly driven by lower perceptual processing. Finally, in Chapter 5, exposure to adversity in adulthood is not associated with either working memory updating or working memory capacity.

**Interpreting task-specific associations with adversity exposure**

The finding that adversity exposure is not associated with specific EF abilities is striking given that lower raw performance on EF tasks is often interpreted as such. Such conclusions are often based on performance on a single task. For instance, lower performance on inhibition tasks has been interpreted as lower inhibition ability (Farah et al., 2006; Fields et al., 2021; Mezzacappa, 2004; Mittal et al., 2015; Noble et al., 2005). Similarly, higher performance on attention-shifting tasks has been interpreted as an enhanced attention shifting ability (Fields et al., 2021; Howard et al., 2020; Mittal et al., 2015; Nweze et al., 2021; Young et al., 2022). My dissertation highlights a crucial limitation of this approach. Task-general processes make it difficult to infer specific abilities based on the performance on a single task, even when using DDM rather than raw performance measures. Even after accounting for task-general processes, though, remaining variance may not capture specific EF abilities (but rather other factors, such as content or familiarity), as suggested by Chapter 3 as well as prior literature (Frischkorn et al., 2019; Löffler et al., 2024).

One reason for not finding ability-specific associations could be that content effects mask the effects of specific EF abilities. Task performance is known to vary with task content (e.g., numbers, letters, or geometric shapes), and studies in cognitive psychology often account for this by sampling tasks with different types of content (Lerche et al., 2020). Some research shows that people from adversity may be particularly sensitive to content effects, and that their performance on EF tasks could be improved by using more real-world content (Young et al., 2022). With the exception of Chapter 2, the studies in this dissertation involved more abstract content, which may be one explanation for lower task-general processing speed. In addition, it may also explain the negative associations in Chapter 3 between childhood adversity and task-specific drift rates on attention-shifting tasks, despite drift rates between tasks correlating weakly. All attention-shifting tasks used abstract content, but the specific type of content differed across tasks. Thus, content effects may have lowered processing on these tasks in specific ways unrelated to the EF ability—the actual target of measurement in these tasks.

### Low reliability of traditional EF tasks

Further down the psychometric path, the elephant in the room is that commonly used EF tasks may not be sufficiently reliable to detect individual differences in EF (Draheim et al., 2019; Hedge et al., 2018; Rouder & Haaf, 2019). Many EF tasks, like the Flanker or Simon task, were developed by experimental psychologists with the aim to obtain robust group-level experimental effects (e.g., the Flanker effect, in which people are on average slower on incongruent trials compared to congruent trials) (Cronbach, 1957). These tasks achieve this by minimizing within-person variability. However, low within-person variability makes them less suitable for studying individual differences. In fact, a recent study showed that over 1,000 trials are needed to obtain reliable estimates of individual differences in the Stroop or Flanker effect (Lee et al., 2023). Needless to say, the studies reported in this dissertation did not even get close to these trial numbers. Nor do the majority of studies in the broader adversity and developmental literature. This is exemplified by the ABCD study, which is currently used in over 1,200 articles (https://abcdstudy.org/publications/), including the study in Chapter 2. The EF tasks included in the ABCD study contain as few as 20 across conditions for the Flanker task.

An inconvenient but important conclusion is that most research in the adversity literature lacks reliable measurements to adequately assess specific EF abilities. In light of this issue, some have argued that large-scale developmental data collections should make fundamentally different trade-offs by lowering the number of participants and increasing the number of trials for cognitive tasks (Lee et al., 2023). Although I agree in theory, there are important constraints that make this unfeasible in practice. In most large cohort studies like the ABCD study, cognitive assessments are only a relatively small part, and so the time spend on cognitive tasks trades off with other important measurements. Even disregarding time as a limiting factor, administering hundreds of trials could decrease motivation and effort. These limitations may be especially large when testing children or people from disadvantaged backgrounds (Niebaum & Munakata, 2023). The statistical techniques used in my dissertation do not by themselves solve this issue. Promising potential solutions include tasks that require accurate but not speeded responding (Draheim et al., 2021; Draheim et al., 2022) and tasks involving gamification, which produce high engagement and arousal, increasing within-person variability (Kucina et al., 2023).

## 6.4 Key finding 3: Adversity exposure is associated with the use of different strategies

My dissertation finds some evidence that exposure to adversity is associated with the use of different cognitive strategies. First, I find evidence for differences in *speed-accuracy trade-offs*, reflecting a person's response caution. A person with higher response caution uses the strategy (deliberate or not) of slowing down their responses in order to increase their accuracy. In Chapter 2, I find that children who experienced more household threat in the preceding year (but not material deprivation) respond more cautiously than children with less exposure to household threat. However, I do not observe differences in response caution in young adults with more exposure to childhood threat and unpredictability (Chapter 4), nor in adults with more exposure to threat and deprivation in childhood or adulthood (Chapter 3). Thus, although exposure to threat is associated with children prioritizing accuracy over speed, the same is not true for (young) adults. Second, Chapter 4 suggests that young adults with more exposure to childhood threat and unpredictability have a more *holistic processing style*, rather than a detail-oriented processing style.

**Speed-accuracy trade-offs**

The results in Chapter 2 are consistent with research on optimal speed-accuracy trade-offs in the face of threats. Individuals across different species tend to be more cautious if they were recently exposed to sources of threat such as violence or predation (Chittka et al., 2009). Making a mistake (e.g., wrongly assuming that there is no predator nearby) can be costly, and therefore it pays to accumulate more information if past environments tended to be more dangerous. However, the opposite is true in the face of immediate danger. In such cases, responding quickly can prevent serious harm, which, all else being equal, outweighs the potential cost of acting too fast (e.g., failing to seize potential resources) (Pirrone et al., 2014).

There is strong evidence that detecting and responding to threat in both scenarios is facilitated by distinct neural pathways: a fast but less accurate pathway in the case of an immediate threat, and a slower but more accurate pathway when there is no immediate threat (LeDoux, 2000). The first relies on short subcortical pathways that provide rapid but coarse information, and do not involve extensive evidence accumulation. Under conditions of stress, people use simpler and faster stimulus-response learning strategies and rely more on habits (Schwabe et al., 2007; Schwabe & Wolf, 2009). In contrast, in the absence of immediate threat, processing relies on longer cortical pathways that do involve evidence accumulation, as modeled using the DDM (Trimmer et al., 2008). Chapter 2 is consistent with this theory: the test setting did not convey an immediate threat and so did not require an immediate response, but children with more exposure to threat did accumulate more evidence.

The results of Chapter 3 and 4 are not consistent with this theory, which may be explained by the temporal gap between the exposure to adversity and the testing session. In Chapter 2, this gap was relatively small; children reported on their exposure to threat in the preceding year. Hence, it is possible that their strategies were still attuned to these recent experiences, which would explain their increased response caution. In Chapter 3 and 4, involving (young) adults, the gap was larger, especially when they retroactively reported on exposure to childhood adversity. Even though Chapter 3 did focus on adversity exposure in adulthood, the adversity measures spanned several years and thus did not necessarily reflect recent adversity exposure. It is possible that differences in how people make speed-accuracy trade-offs in response to threat exposure remain plastic, such that a preference for accuracy over speed may diminish or even disappear if threats become less frequent. This is an open and interesting question for future research.

### Holistic versus detail-oriented processing style

Beyond speed-accuracy trade-offs, Chapter 4 also provides evidence for differences in how people with more exposure to childhood adversity process information. Across three studies, the strength of perceptual processing on the Flanker task was lower for people with more exposure to both violence and unpredictability, which may indicate a deficit in information processing. However, the strength of perceptual processing interacted with a person's processing style. In the context of the Flanker task, strength of perceptual processing refers to the amount of visual information that people extract from the arrows. For people with more exposure to childhood violence (and to a lesser extent unpredictability), lower strength of perceptual processing was related to more holistic processing rather than featural or detail-oriented processing. In contrast, people with less exposure to childhood violence had a higher strength of perceptual processing, which was related to more detail-oriented processing.

Although the interaction between perceptual processing and holistic processing requires more research, I speculate that a more holistic processing style in people with more adversity exposure may (partially) account for lower strength of perceptual processing. It could relate to the speed-accuracy trade-off discussed above: Aside from taking more time to accumulate information, adopting a more holistic processing style facilitates the detection of potential threats compared to a more focused processing style. The Shrinking Spotlight Model used to decompose Flanker performance in Chapter 4 distinguishes between a processing parameter (i.e., strength of perceptual processing) and two attention parameters (i.e., the initial width of the attention scope, and the rate at which attention narrows over time). A more holistic processing style could affect both. On the one hand, holistic processing could lower the strength of perceptual processing as stimuli are processed as a whole instead of as individual sources of information. On the other hand, attention would be spread out more evenly across all stimuli, and narrowing attention down to the central target would be more difficult.

Unfortunately though, I could not accurately recover the two attention parameters in isolation, and instead computed a ratio between them (in line with White et al., 2018). Future studies with a larger number of trials may be able to recover the attention parameters. Additionally, future research could include more direct measures of attention such as eye-tracking and pupillometry. Previous research suggests that people with a more holistic processing style have fewer fixations on individual items as well as larger saccades (Schreiter & Vogel, 2023, 2024). Thus, it would be insightful to investigate whether such attention features provide a common explanation for holistic processing as well as a lower strength of perceptual processing on inhibition tasks.

## 6.5 Developing a roadmap for adversity research

### Integrating deficit and adaptation frameworks

Adversity researchers generally acknowledge that exposure to adversity can both impair and lead to adaptations in cognitive abilities (Ellis et al., 2022; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013; Noble et al., 2021). Several studies suggest that the same person can show deficits in some abilities yet enhancements in other abilities.
For instance, people with more exposure to adversity were slower on inhibition tasks but faster on attention-shifting tasks (Fields et al., 2021; Mittal et al., 2015), and slower on a working memory capacity task but faster on a working memory updating task (Young et al., 2018). As has become clear throughout my dissertation, such comparisons of individual tasks are problematic given that tasks share cognitive processes. A more realistic vantage point appears to be that both types of processes can operate within the same task. Using cognitive modeling, future research will be well-positioned to test more precise predictions about how deficit and adaptation processes interact.

Researchers need to deal with the fact that task performance is influenced by both task-general and ability-specific processes. Deficit frameworks predict impairments in both specific abilities as well as general processing (e.g., associated with impairments in more localized as well as more widely connected brain networks) (Sheridan & McLaughlin, 2014; Tucker-Drob, 2013). Still, as impairments in general and specific processes may have different origins (e.g., functional or structural changes in the brain), differentiating them using cognitive modeling and structural equation modeling affords testing more precise predictions. Arguably, the existence of task-general processes poses a bigger challenge for adaptation frameworks, which predict that specific types of adversity enhance specific cognitive abilities (Ellis et al., 2022; Frankenhuis, Young, et al., 2020; Frankenhuis & Weerth, 2013). Testing such predictions will require accounting for general processes and ideally sampling two or more tasks for each ability.

Abilities enhanced by adversity may even cross the boundaries of traditional EF tasks. For instance, several studies have found that people from lower socioeconomic backgrounds are more attentive to task-irrelevant sounds (D'angiulli, Van Roon, et al., 2012; Giuliano et al., 2018; Hao & Hu, 2024; Stevens et al., 2009). Children from lower socioeconomic backgrounds also appear more attentive to peripheral visual information (Mezzacappa, 2004). Similarly, exposure to adversity might lead to a more diffuse scope of attention to facilitate tracking the environment for potential threats and opportunities. We did not find support for our initial hypothesis (see the Introduction of Chapter 4) that this might make people better at detecting subtle changes and peripheral stimuli, i.e., an enhanced ability to detect specific stimuli in the broader environment. However, as discussed in section 6.4, we did find a tendency towards holistic processing. This may be an alternative manifestation of diffuse attention, where people do not so much attend to individual features in the periphery, but rather do so more holistically. Both cognitive modeling and structural equation modeling can help to illuminate such phenotypes and how they affect performance across traditional EF tasks.

Integrating deficit and adaptation frameworks also requires quantifying support in favor of the null hypothesis (i.e., intact ability), rather than only against the null hypothesis (i.e., impaired or enhanced ability) (Harms & Lakens, 2018; Lakens et al., 2018). Cognitive adaptations may not always lead to enhancements, but could also translate to intact ability, especially when performance is simultaneously influenced by deficits (Bignardi et al., 2024; Young et al., 2024). Using practical equivalence testing, I find some evidence for intact specific abilities after accounting for task-general processing speed, especially in Chapter 2. Equally importantly, in many cases I found inconclusive results, with evidence supporting neither adversity-related differences nor practical equivalence. Throughout, I have used a standardized effect of 0.1 as the cut-off for practical equivalence, with effects smaller than 0.1 considered practically equivalent to zero. This cut-off is arbitrary: some small effects can have a substantial impact on the population level, and conversely, some effects above 0.1 may not be all that meaningful. As researchers learn more about which effects sizes are associated with meaningful outcomes (and which are not), they can adopt more theory-guided cut-offs.

**Better understanding content and context effects on EF performance**

Some developmental psychologists argue that abstract EF tasks may disadvantage people from more disadvantaged backgrounds, e.g., due to less formal education (Doebel, 2020; Frankenhuis, Young, et al., 2020; Miller-Cotto et al., 2022; Niebaum & Munakata, 2023). Common EF tasks may disadvantage children from minority groups because they were developed for, and normed based on children from majority groups (Miller-Cotto et al., 2022). Children from minority groups may in part perform lower because EF tasks are divorced from their everyday experiences and cultural and social norms. They involve unfamiliar researchers and test settings that are unlike the environments they are used to (Doebel, 2020). From an adaptive perspective, it has been argued that people may perform best when task conditions, including the stimuli that are used, match the conditions in which they developed their cognitive abilities (Frankenhuis, Young, et al., 2020). Consistent with this idea, real-world content has been found to affect performance on EF tasks, and in some cases this effect is larger for people with more exposure to adversity (Young et al., 2022). Finally, abstract testing conditions may even lower children's willingness to *engage* EF, for instance, because the task does not seem relevant or because it does not seem worth the effort (Niebaum & Munakata, 2023). Thus, to understand the effect of adversity on EF performance, we may need to understand performance in people's broader ecological context.

Although it is important to develop more equitable and valid EF tasks, such tasks risk the same psychometric limitations that have been central in my dissertation. Researchers should not assume that more ecologically relevant tasks are less susceptible to the influence of general processes and speed-accuracy trade-offs. For instance, different types of content could affect performance through different pathways: it may influence general processes, specific abilities, response caution, or a combination of these and other factors. Cognitive modeling and structural equation modeling can play a key role in better understanding which cognitive processes are affected by different types of task manipulations.

To focus on one example, cognitive modeling can illuminate which dimensions of stimulus content are responsible for closing (or widening) performance gaps. One key dimension may be how *familiar* the content is to the person taking the test, that is, the extent to which a stimulus has been encountered before (Niebaum & Munakata, 2023). On the one hand, more familiar task content may increase ability-specific drift rates. For instance, inhibiting distractors may be easier when the target stimulus is more familiar, and keeping track of information in working memory may be easier if the information relates to previous experiences. On the other hand, more familiar task content may increase task-general drift rate, for instance, if familiar content reduces the cognitive burden of the task regardless of the EF ability that is targeted. Performance differences may also arise from other content dimensions, and their influence on performance could stem from other cognitive processes. Stimuli that are more *valenced* could influence response caution (e.g., being more careful when a stimulus makes you anxious) or, in some cases, response bias (e.g., a bias towards threatening stimuli). Finally, real-world stimuli may often be more visually *complex* than standard abstract stimuli (e.g., numbers, shapes). This could make it more difficult to visually encode the stimulus, increasing non-decision times. Testing these effects using cognitive modeling can illuminate if and why certain types of content negatively or positively affect performance.

## 6.6 Concluding remarks

*"We pass through this world but once. Few tragedies can be more extensive than the stunting of life, few injustices deeper than the denial of an opportunity to strive or even to hope, by a limit imposed from without, but falsely identified as lying within."* Stephen J. Gould (1980). *The mismeasure of man*.

Cognitive assessments affect millions of lives each year. Performance scores influence academic trajectories, selection of people into jobs, and are at the basis for a variety of interventions and policies. They also shape how people view their own potential, and how their potential is viewed by others. It is therefore crucial that our interpretations of cognitive performance accurately reflect a person's ability. My dissertation shows that for people from adverse environments, who tend to perform lower on cognitive tasks, this may often not be so. Hence, research may underestimate a person's true EF abilities, and attempt to fix things that are not actually 'broken', while potentially overlooking areas requiring attention. Fortunately, adversity researchers can stand on the shoulders of decades of research in cognitive psychology that allows for a more precise assessment of cognitive abilities. In particular, cognitive modeling will be an indispensable instrument in the toolbox of the next generation of adversity researchers.

The use of DDM and structural equation modeling need not be limited to basic scientific research; instead, it could be directly used in applied contexts, such as clinical or high-stakes testing. Now that digital testing is widespread and affordable, there is no good reason to hold onto raw performance measures. Instead, screening and assessment batteries could directly incorporate DDM and structural equation modeling to provide more meaningful estimates of cognitive processes. Beyond that initial step, insights from these techniques could be used to tailor assessments to individuals. For instance, based on future scientific insights, assessment batteries could personalize instructions and task content in response to initial estimates of cognitive processes, and track their change over time. This way, cognitive modeling has the potential to directly impact children's and adults' lives.

# Footnotes

[1] It is also possible to have the decision boundaries correspond to distinct response options. For instance, if the task requires people to classify faces as expressing either an angry or happy emotion, the boundaries could correspond to the 'angry' and 'happy' response, respectively. This specification is useful if the research question pertains to decision preferences (e.g., do people tend to prefer option A over option B? Do people with more exposure to threat tend to interpret facial expressions as more negative?).

[2] There are two caveats to this statement. First, the standard DDM also contains the starting point of the evidence accumulation process and provides a measure of response bias. When the process starts closer to one boundary relative to the other, it reaches this boundary faster and more frequently (also increasing the false positive rate), while responses terminating at the other boundary are slower and less frequent. Modeling the starting point makes most sense if the decision boundaries correspond to distinct response options (e.g., angry versus happy face) rather than correct versus incorrect responses. In the latter case, the response bias parameter is usually fixed to be equidistant to each boundary. As this is the case throughout this dissertation, I do not consider the starting point here. Second, the DDM also allows for additional parameters that capture inter-trial variability in drift rate, boundary separation, and/or non-decision time. For instance, drift rates may decrease as people start to lose motivation. Simulation studies indicate that several hundreds of trials are necessary to obtain stable estimates of these variability parameters, many more than were used in the cognitive tasks included in this dissertation.

[3] A fourth DDM parameter, the starting point, represents an initial bias towards one of the two decision options (e.g., a tendency to classify facial expressions as angry that extends to neutral faces). Note that allowing the starting point to vary only makes sense if response options differ in valence (e.g., happy and angry faces, which the current study does not include and thus is unable to examine).

[4] The preregistration also included the Picture Vocabulary Task. However, after accessing the data we realized that this task was implemented using computerized adaptive testing (Luciana et al., 2018). This makes it unsuitable for DDM, as the model assumes the level of difficulty is the same across trials.

# Bibliography


Ahmed, S., Tang, S., Waters, N., & Davis-Kean, P. (2018). Executive function and academic achievement: Longitudinal relations from early childhood to adolescence. *Journal of Educational Psychology*, *111*. https://doi.org/10.1037/edu0000296

Algarin, C., Karunakaran, K. D., Reyes, S., Morales, C., Lozoff, B., Peirano, P., & Biswal, B. (2017). Differences on brain connectivity in adulthood are present in subjects with iron deficiency anemia in infancy. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10.3389/fnagi.2017.00054

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Ankan, A., Wortel, I. M. N., & Textor, J. (2021). Testing graphical causal models using the R package “dagitty.” *Current Protocols*, *1*(2), e45. https://doi.org/10.1002/cpz1.45

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological Science*, *15*(9), 610–615. https://doi.org/10.1111/j.0956-7976.2004.00728.x

Assari, S. (2020). Mental rotation in American children: Diminished returns of parental education in black families. *Pediatric Reports*, *12*(3), 130–141. https://doi.org/10.3390/pediatric12030028

Barkley, R. A. (2012). *Executive functions: What they are, how they work, and why they evolved*. Guilford Press.

Bastian, C. C. von, Blais, C., Brewer, G., Gyurkovics, M., Hedge, C., Kałamała, P., Meier, M., Oberauer, K., Rey-Mermet, A., Rouder, J. N., Souza, A. S., Bartsch, L. M., Conway, A. R. A., Draheim, C., Engle, R. W., Friedman, N. P., Frischkorn, G. T., Gustavson, D. E., Koch, I., … Wiemers, E. (2020). *Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations*. PsyArXiv. https://doi.org/10.31234/osf.io/x3b9k

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, *22*, 507–526. https://doi.org/10.1037/met0000077

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/10.1037/a0024454

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bignardi, G., Mareva, S., & Astle, D. E. (2024). Parental socioeconomic status weakly predicts specific cognitive and academic skills beyond general cognitive ability. *Developmental Science*, *27*(2), e13451. https://doi.org/10.1111/desc.13451

Bij, A. K. van der, Weerd, S. de, Cikot, R. J. L. M., Steegers, E. A. P., & Braspenning, J. C. C. (2003). Validation of the Dutch Short Form of the State Scale of the Spielberger State-Trait Anxiety Inventory: Considerations for usage in screening outcomes. *Community Genetics*, *6*(2), 84–87. https://doi.org/10.1159/000073003

Blair, C., & Raver, C. C. (2012). Child development in the context of adversity. *The American Psychologist*, *67*(4), 309–318. https://doi.org/10.1037/a0027493

Blair, C., & Raver, C. C. (2014). Closing the achievement gap through modification of neurocognitive and neuroendocrine function: Results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PLOS ONE*, *9*(11), e112393. https://doi.org/10.1371/journal.pone.0112393

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(3-4), 296–312. https://doi.org/10.1111/j.1469-7610.2006.01611.x

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences*, *33*(1), 10–16. https://doi.org/10.1016/j.tins.2009.09.002

Bos, K. J., Fox, N., Zeanah, C. H., & Nelson III, C. A. (2009). Effects of early psychosocial deprivation on the development of memory and executive function. *Frontiers in Behavioral Neuroscience*, *3*, 16. https://doi.org/10.3389/neuro.08.016.2009

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & García Coll, C. (2001). The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status. *Child Development*, *72*(6), 1844–1867. https://doi.org/10.1111/1467-8624.t01-1-00382

Braem, S. (2017). Conditioning task switching behavior. *Cognition*, *166*, 272–276. https://doi.org/10.1016/j.cognition.2017.05.037

Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Daily variability in working memory is coupled with negative affect: The role of attention and motivation. *Emotion*, *12*(3), 605–617. https://doi.org/10.1037/a0024436

Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, *29*(6), 624–630. https://doi.org/10.1177/0963721420969371

Burgoyne, A. P., Mashburn, C. A., Tsukahara, J. S., & Engle, R. W. (2022). Attention control and process overlap theory: Searching for cognitive processes underpinning the positive manifold. *Intelligence*, *91*, 101629. https://doi.org/10.1016/j.intell.2022.101629

Burnham, K. P., & Anderson, D. R. (Eds.). (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer New York.

Bywaters, P., Bunting, L., Davidson, G., Hanratty, J., Mason, W., McCartan, C., & Steils, N. (2016). *The relationship between poverty, child abuse and neglect: An evidence review*. Joseph Rowntree Foundation. https://www.jrf.org.uk/child-poverty/the-relationship-between-poverty-child-abuse-and-neglect-an-evidence-review

Carlozzi, N. E., Beaumont, J. L., Tulsky, D. S., & Gershon, R. C. (2015). The NIH Toolbox Pattern Comparison Processing Speed Test: Normative data. *Archives of Clinical Neuropsychology*, *30*(5), 359–368. https://doi.org/10.1093/arclin/acv031

Cermakova, P., Chlapečka, A., Csajbók, Z., Andrýsková, L., Brázdil, M., & Marečková, K. (2023). Parental education, cognition and functional connectivity of the salience network. *Scientific Reports*, *13*(1), 2761. https://doi.org/10.1038/s41598-023-29508-w

Chambers, C. D., & Tzavella, L. (2021). The past, present and future of Registered Reports. *Nature Human Behaviour*, *6*(1), 29–42. https://doi.org/10.1038/s41562-021-01193-7

Chiappe, P., Hasher, L., & Siegel, L. S. (2000). Working memory, inhibitory control, and reading disability. *Memory & Cognition*, *28*(1), 8–17. https://doi.org/10.3758/BF03211570

Chittka, L., Skorupski, P., & Raine, N. E. (2009). Speed–accuracy tradeoffs in animal decision making. *Trends in Ecology & Evolution*, *24*(7), 400–407. https://doi.org/10.1016/j.tree.2009.02.010

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. https://doi.org/https://doi.org/10.3758/BF03196772

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cribbie, R. A. (2007). Multiplicity control in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(1), 98–112. https://doi.org/10.1080/10705510709336738

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671–684. https://doi.org/10.1037/h0043943

Curran, P. J., McGinley, J. S., Bauer, D. J., Hussong, A. M., Burns, A., Chassin, L., Sher, K., & Zucker, R. (2014). A moderated nonlinear factor model for the development of commensurate measures in integrative data analysis. *Multivariate Behavioral Research*, *49*(3), 214–231. https://doi.org/10.1080/00273171.2014.889594

D'angiulli, A., Lipina, S., & Olesinska, A. (2012). Explicit and implicit issues in the developmental cognitive neuroscience of social inequality. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00254

D'angiulli, A., Van Roon, P. M., Weinberg, J., Oberlander, T., Grunau, R., Hertzman, C., & Maggi, S. (2012). Frontal EEG/ERP correlates of attentional processes, cortisol and motivational states in adolescents from lower and higher socioeconomic status. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00306

Dai, T., Guo, Y., & Alzheimer's Disease Neuroimaging Initiative. (2017). Predicting individual brain functional connectivity using a Bayesian hierarchical model. *NeuroImage*, *147*, 772–787. https://doi.org/10.1016/j.neuroimage.2016.11.048

Daly, M., & Wilson, M. (2005). Carpe diem: Adaptation and devaluing the future. *The Quarterly Review of Biology*, *80*(1), 55–60. https://doi.org/10.1086/431025

Dang, J. (2017). Commentary: The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Frontiers in Psychology*, *8*, 1711. https://doi.org/10.3389/fpsyg.2017.01711

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

DeBruine, L. (2021). *Faux: Simulation for factorial designs*. Zenodo. https://doi.org/10.5281/zenodo.5513951

DeJoseph, M. L., Herzberg, M. P., Sifre, R. D., Berry, D., & Thomas, K. M. (2022). Measurement matters: An individual differences examination of family socioeconomic factors, latent dimensions of children's experiences, and resting state functional brain connectivity in the ABCD sample. *Developmental Cognitive Neuroscience*, *53*, 101043. https://doi.org/10.1016/j.dcn.2021.101043

Del Giudice, M., & Crespi, B. J. (2018). Basic functional trade-offs in cognition: An integrative framework. *Cognition*, *179*, 56–70. https://doi.org/10.1016/j.cognition.2018.06.008

Del Giudice, M., & Gangestad, S. W. (2021). A traveler's guide to the multiverse: Promises, pitfalls, and a framework for the evaluation of analytic decisions. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920954925

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics*, *1*(3), 111–134. https://doi.org/10.1257/app.1.3.111

Denwood, M. J. (2016). Runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software*, *71*, 1–25. https://doi.org/10.18637/jss.v071.i09

Diemer, M. A., Mistry, R. S., Wadsworth, M. E., López, I., & Reimers, F. (2013). Best practices in conceptualizing and measuring social class in psychological research. *Analyses of Social Issues and Public Policy*, *13*(1), 77–113. https://doi.org/10.1111/asap.12001

Distefano, R., Fiat, A. E., Merrick, J. S., Slotkin, J., Zelazo, P. D., Carlson, S. M., & Masten, A. S. (2021). NIH Toolbox executive function measures with developmental extensions: Reliability and validity with preschoolers in emergency housing. *Child Neuropsychology*, *27*(6), 709–717. https://doi.org/10.1080/09297049.2021.1888905

Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science*, *15*(4), 942–956. https://doi.org/10.1177/1745691620904771

Donders, F. C. (1869). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431. https://doi.org/10.1016/0001-6918(69)90065-1

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. https://doi.org/10.1037/bul0000192

Draheim, C., Pak, R., Draheim, A. A., & Engle, R. W. (2022). The role of attention control in complex real-world tasks. *Psychonomic Bulletin & Review*, *29*. https://doi.org/10.3758/s13423-021-02052-2

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*, *150*(2), 242–275. https://doi.org/10.1037/xge0000783

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2017). Moving beyond correlations in assessing the consequences of poverty. *Annual Review of Psychology*, *68*(1), 413–434. https://doi.org/10.1146/annurev-psych-010416-044224

Duncan, G. J., Ziol-Guest, K. M., & Kalil, A. (2010). Early-childhood poverty and adult attainment, behavior, and health. *Child Development*, *81*(1), 306–325. https://doi.org/10.1111/j.1467-8624.2009.01396.x

Duquennois, C. (2022). Fictional money, real costs: Impacts of financial salience on disadvantaged students. *American Economic Review*, *112*(3), 798–826. https://doi.org/10.1257/aer.20201661

Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Weissberg, R. P., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Duval, E. R., Garfinkel, S. N., Swain, J. E., Evans, G. W., Blackburn, E. K., Angstadt, M., Sripada, C. S., & Liberzon, I. (2017). Childhood poverty is associated with altered hippocampal function and visuospatial memory in adulthood. *Developmental Cognitive Neuroscience*, *23*, 39–44. https://doi.org/10.1016/j.dcn.2016.11.006

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 170–189. https://doi.org/10.1037/a0017891

Ellis, B. J., Abrams, L., Masten, A., Sternberg, R., Tottenham, N., & Frankenhuis, W. (2022). Hidden talents in harsh environments. *Development and Psychopathology*, *34*, 95–113. https://doi.org/10.1017/S0954579420000887

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/10.1177/1745691617693054

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk. *Human Nature*, *20*(2), 204–268. https://doi.org/10.1007/s12110-009-9063-7

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*(4), 225–240. https://doi.org/10.3758/BF03211502

Evans, G. W., & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences*, *106*(16), 6545–6549. https://doi.org/10.1073/pnas.0811910106

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/10.1016/j.brainres.2006.06.072

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/10.1111/desc.13133

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Re-*

*view of Psychology*, *67*(1), 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

Francis, E. R., Tsaligopoulou, A., Stock, S. E., Pingault, J., & Baldwin, J. R. (2023). Subjective and objective experiences of childhood adversity: A meta-analysis of their agreement and relationships with psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *64*(8), 1185–1199. https://doi.org/10.1111/jcpp.13803

Frankenhuis, W. E., & Bijlstra, G. (2018). Does exposure to hostile environments predict enhanced emotion detection? *Collabra: Psychology*, *4*(1), 18. https://doi.org/10.1525/collabra.127

Frankenhuis, W. E., Panchanathan, K., & Nettle, D. (2016). Cognition in harsh and unpredictable environments. *Current Opinion in Psychology*, *7*, 76–80. https://doi.org/10.1016/j.copsyc.2015.08.011

Frankenhuis, W. E., Vries, S. A. de, 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/10.1111/desc.12835

Frankenhuis, W. E., & Weerth, C. de. (2013). Does early-life exposure to stress shape or impair cognition? *Current Directions in Psychological Science*, *22*(5), 407–412. https://doi.org/10.1177/0963721413484324

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

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology. General*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Frischkorn, G. T., Bastian, C. C. von, Souza, A. S., & Oberauer, K. (2022). Individual differences in updating are not related to reasoning ability and working memory capacity. *Journal of Experimental Psychology: General*, *151*(6), 1341–1357. https://doi.org/10.1037/xge0001141

Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2019). Processing speed, working memory, and executive functions: Independent or inter-related predictors of general intelligence. *Intelligence*, *75*, 95–110. https://doi.org/10.1016/j.intell.2019.05.003

Frost, A., Moussaoui, S., Kaur, J., Aziz, S., Fukuda, K., & Niemeier, M. (2021). Is the n-back task a measure of unstructured working memory capacity? Towards understanding its connection to other working memory tasks. *Acta Psychologica*, *219*, 103398. https://doi.org/10.1016/j.actpsy.2021.103398

Fuhrmann, D., Simpson-Kent, I. L., Bathelt, J., The CALM Team, & Kievit, R. A. (2020). A hierarchical watershed model of fluid intelligence in childhood and adolescence. *Cerebral Cortex*, *30*(1), 339–352. https://doi.org/10.1093/cercor/bhz091

Ganschow, B., Zebel, S., Van Der Schalk, J., Hershfield, H. E., & Van Gelder, J.-L. (2023). Adolescent stressful life events predict future self- connectedness in adulthood. *The Journal of Early Adolescence*, *44*(9). https://doi.org/10.1177/02724316231216380

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., & Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, *32*, 16–22. https://doi.org/10.1016/j.dcn.2018.04.004

Gellci, K., Marusak, H. A., Peters, C., Elrahal, F., Iadipaolo, A. S., & Rabinak, C. A. (2019). Community and household-level socioeconomic disadvantage and functional organization of the salience and emotion network in children and adolescents. *NeuroImage*, *184*, 729–740. https://doi.org/10.1016/j.neuroimage.2018.09.077

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457–472. https://doi.org/10.1214/ss/1177011136

Giuliano, R. J., Karns, C. M., Roos, L. E., Bell, T. A., Petersen, S., Skowron, E. A., Neville, H. J., & Pakulak, E. (2018). Effects of early adversity on neural mechanisms of distractor suppression are mediated by sympathetic nervous system activity in preschool-aged children. *Developmental Psychology*, *54*(9), 1674–1686. https://doi.org/10.1037/dev0000499

Glynn, L. M., Stern, H. S., Howland, M. A., Risbrough, V. B., Baker, D. G., Nievergelt, C. M., Baram, T. Z., & Davis, E. P. (2019). Measuring novel antecedents of mental illness: The Questionnaire of Unpredictability in Childhood. *Neuropsychopharmacology*, *44*(5), 876–882. https://doi.org/10.1038/s41386-018-0280-9

Goltermann, J., Meinert, S., Hülsmann, C., Dohm, K., Grotegerd, D., Redlich, R., Waltemate, L., Lemke, H., Thiel, K., Mehler, D. M. A., Enneking, V., Borgers, T., Repple, J., Gruber, M., Winter, N., Hahn, T., Brosch, K., Meller, T., Ringwald, K. G., … Dannlowski, U. (2023). Temporal stability and state-dependence of retrospective self-reports of childhood maltreatment in healthy and depressed adults. *Psychological Assessment*, *35*(1), 12–22. https://doi.org/10.1037/pas0001175

Goodman, J. B., Freeman, E. E., & Chalmers, K. A. (2019). The relationship between early life stress and working memory in adulthood: A systematic review and meta-analysis. *Memory*, *27*(6), 868–880. https://doi.org/10.1080/09658211.2018.1561897

Grange, J. A. (2016). Flankr: An R package implementing computational models of attentional selectivity. *Behavior Research Methods*, *48*(2), 528–541. https://doi.org/10.3758/s13428-015-0615-y

Grange, J. A., & Rydon-Grange, M. (2022). Computational modelling of attentional selectivity in depression reveals perceptual deficits. *Psychological Medicine*, *52*(5), 904–913. https://doi.org/10.1017/S0033291720002652

Gruszka, A., & Nęcka, E. (2017). Limitations of working memory capacity: The cognitive and social consequences. *European Management Journal*, *35*(6), 776–784. https://doi.org/10.1016/j.emj.2017.07.001

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, *16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Guo, Z., Zou, J., He, C., Tan, X., Chen, C., & Feng, G. (2020). The Importance of cognitive and mental factors on prediction of job performance in chinese high-speed railway dispatchers. *Journal of Advanced Transportation*, *2020*, e7153972. https://doi.org/10.1155/2020/7153972

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/10.1038/nrn2897

Hakim, N., Simons, D. J., Zhao, H., & Wan, X. (2017). Do Easterners and Westerners differ in visual cognition? A preregistered examination of three visual cognition tasks. *Social Psychological and Personality Science*, *8*(2), 142–152. https://doi.org/10.1177/1948550616667613

Hanson, J. L., Chung, M. K., Avants, B. B., Rudolph, K. D., Shirtcliff, E. A., Gee, J. C., Davidson, R. J., & Pollak, S. D. (2012). Structural variations in prefrontal cortex mediate the relationship between early childhood stress and spatial working memory. *Journal of Neuroscience*, *32*(23), 7917–7925. https://doi.org/10.1523/JNEUROSCI.0307-12.2012

Hao, Y., & Hu, L. (2024). Lower childhood socioeconomic status Is associated with greater neural responses to ambient auditory changes in adulthood. *Journal of Cognitive Neuroscience*, *36*(6), 979–996. https://doi.org/10.1162/jocn_a_02151

Harms, C., & Lakens, D. (2018). Making ’null effects’ informative: Statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research*, *3*(Suppl 2), 382–393. https://doi.org/10.18053/jctres.03.2017S2.007

Hazel, N. A., Hammen, C., Brennan, P. A., & Najman, J. (2008). Early childhood adversity and adolescent depression: The mediating role of continued stress. *Psychological Medicine*, *38*(4), 581–589. https://doi.org/10.1017/S0033291708002857

Hedge, C., Powell, G., Bompas, A., & Sumner, P. (2022). Strategy and processing speed eclipse individual differences in control ability in conflict tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(10), 1448–1469. https://doi.org/10.1037/xlm0001028

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied survey data analysis*. Chapman & Hall.

Hilger, K., Ekman, M., Fiebach, C. J., & Basten, U. (2017). Efficient hubs in the intelligent brain: Nodal efficiency of hub regions in the salience network is associated with general intelligence. *Intelligence*, *60*, 10–25. https://doi.org/10.1016/j.intell.2016.11.001

Hofmarcher, T. (2021). The effect of education on poverty: A European perspective. *Economics of Education Review*, *83*, 102124. https://doi.org/10.1016/j.econedurev.2021.102124

Hoijtink, H., Mulder, J., Lissa, C. van, & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*, *24*(5), 539–556. https://doi.org/10.1037/met0000201

Hostinar, C. E., & Gunnar, M. R. (2013). The developmental effects of early life stress: An overview of current theoretical frameworks. *Current Directions in Psychological Science*, *22*(5), 400–406. https://doi.org/10.1177/0963721413488889

Howard, S. J., Cook, C. J., Everts, L., Melhuish, E., Scerif, G., Norris, S., Twine, R., Kahn, K., & Draper, C. E. (2020). Challenging socioeconomic status: A cross-cultural comparison of early executive function. *Developmental Science*, *23*(1), e12854. https://doi.org/10.1111/desc.12854

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological Review*, *117*(3), 759–784. https://doi.org/10.1037/a0019471

Huizinga, M., Burack, J. A., & Van der Molen, M. W. (2010). Age-related change in shifting attention between global and local levels of hierarchical stimuli. *Journal of Cognition and Development*, *11*(4), 408–436. https://doi.org/10.1080/15248371003700031

Johnson, D. J., Hopwood, C. J., Cesario, J., & Pleskac, T. J. (2017). Advancing research on cognitive processes in social and personality psychology: A hierarchical Drift Diffusion Model primer. *Social Psychological and Personality Science*, *8*(4), 413–423. https://doi.org/10.1177/1948550617703174

Johnson, D., Policelli, J., Li, M., Dharamsi, A., Hu, Q., Sheridan, M. A., McLaughlin, K. A., & Wade, M. (2021). Associations of early-life threat and deprivation with executive functioning in childhood and adolescence: A systematic review and meta-analysis. *JAMA Pediatrics*, *175*(11), e212511. https://doi.org/10.1001/jamapediatrics.2021.2511

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Karr, J., Areshenkoff, C., Rast, P., Hofer, S., Iverson, G., & Garcia-Barrera, M. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*. https://doi.org/10.1037/bul0000160

Key, N., Prager, D., & Burns, C. (2017). *Farm household income volatility: An analysis using panel data from a national survey*. US Department of Agriculture, Economic Research Service. https://www.ers.usda.gov/publications/pub-details/?pubid=82563

Kievit, R. A., Davis, S. W., Griffiths, J., Correia, M. M., Cam-CAN, & Henson, R. N. (2016). A watershed model of individual differences in fluid intelligence. *Neuropsychologia*, *91*, 186–198. https://doi.org/10.1016/j.neuropsychologia.2016.08.008

Kievit, R. A., McCormick, E. M., Fuhrmann, D., Deserno, M. K., & Orben, A. (2022). Using large, publicly available data sets to study adolescent development: Opportunities and challenges. *Current Opinion in Psychology*, *44*, 303–308. https://doi.org/10.1016/j.copsyc.2021.10.003

Klein, K., & Boals, A. (2001). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology*, *15*(5), 565–579. https://doi.org/10.1002/acp.727

Kovacs, K., & Conway, A. R. A. (2016). Process Overlap Theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, *27*(3), 151–177. https://doi.org/10.1080/1047840X.2016.1153946

Kovacs, K., & Conway, A. R. A. (2019). A unified cognitive/differential approach to human intelligence: Implications for IQ testing. *Journal of Applied Research in Memory and Cognition*, *8*(3), 255–272. https://doi.org/10.1016/j.jarmac.2019.05.003

Kretzschmar, A., & Gignac, G. E. (2019). At what sample size do latent variable correlations stabilize? *Journal of Research in Personality*, *80*, 17–22. https://doi.org/10.1016/j.jrp.2019.03.007

Kübel, S. L., Deitzer, J. R., Frankenhuis, W. E., Ribeaud, D., Eisner, M. P., & Gelder, J.-L. van. (2023). The shortsighted victim: Short-term mindsets mediate the link between victimization and later offending. *Journal of Criminal Justice*, *86*, 102062. https://doi.org/10.1016/j.jcrimjus.2023.102062

Kucina, T., Wells, L., Lewis, I., Salas, K. de, Kohl, A., Palmer, M. A., Sauer, J. D., Matzke, D., Aidman, E., & Heathcote, A. (2023). Calibration of cognitive tests to address the reliability paradox for decision-conflict tasks. *Nature Communications*, *14*(1), 2234. https://doi.org/10.1038/s41467-023-37777-2

Lacey, R. E., Howe, L. D., Kelly-Irving, M., Bartley, M., & Kelly, Y. (2022). The clustering of adverse childhood experiences in the Avon longitudinal study of parents and children: Are gender and poverty important? *Journal of Interpersonal Violence*, *37*(5-6), 2218–2241. https://doi.org/10.1177/0886260520935096

Lakens, D. (2024). When and how to deviate From a preregistration. *Collabra: Psychology*, *10*(1), 117094. https://doi.org/10.1525/collabra.117094

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

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*(1), 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

Lee, H. J., Smith, D. M., Hauenstein, C., Dworetsky, A., Kraus, B. T., Dorn, M., Nee, D., & Gratton, C. (2023). *Precise individual measures of inhibitory control*. OSF. https://doi.org/10.31234/osf.io/rj2bu

Lerche, V., Krause, M. von, Voss, A., Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2020). Diffusion modeling and intelligence: Drift rates show both domain-general and domain-specific relations with intelligence. *Journal of Experimental Psychology: General*, *149*(12), 2207–2249. https://doi.org/10.1037/xge0000774

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, *49*(2), 513–537. https://doi.org/10.3758/s13428-016-0740-2

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. SAGE publications.

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/10.1037/dev0000601

Liu, S., Zalewski, M., Lengua, L., Gunnar, M. R., Giuliani, N., & Fisher, P. A. (2022). Material hardship level and unpredictability in relation to U.S. Households' family interactions and emotional well-being: Insights from the COVID-19 pandemic. *Social Science & Medicine*, *307*, 115173. https://doi.org/10.1016/j.socscimed.2022.115173

Löffler, C., Frischkorn, G. T., Hagemann, D., Sadus, K., & Schubert, A.-L. (2024). The common factor of executive functions measures nothing but speed of information uptake. *Psychological Research*, *88*, 1092–1114. https://doi.org/10.1007/s00426-023-01924-7

Luciana, M., Bjork, J. M., Nagel, B. J., Barch, D. M., Gonzalez, R., Nixon, S. J., & Banich, M. T. (2018). Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Developmental Cognitive Neuroscience*, *32*, 67–79. https://doi.org/10.1016/j.dcn.2018.02.006

Lurie, L. A., Rosen, M. L., Weissman, D. G., Machlin, L., Lengua, L., Sheridan, M. A., & McLaughlin, K. A. (2024). Cognitive stimulation as a mechanism linking socioeconomic status and neural function supporting working memory: A longitudinal fMRI study. *Cerebral Cortex*, *34*(2), bhad545. https://doi.org/10.1093/cercor/bhad545

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *The British Journal of Clinical Psychology*, *31*(3), 301–306. https://doi.org/10.1111/j.2044-8260.1992.tb00997.x

Mashburn, C. A., Barnett, M. K., & Engle, R. W. (2023). Processing speed and executive attention as causes of intelligence. *Psychological Review*, *131*(3), 664–694. https://doi.org/10.1037/rev0000439

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, *56*(3), 227–238. https://doi.org/10.1037/0003-066X.56.3.227

Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the Confusion, Hubbub, and Order Scale. *Journal of Applied Developmental Psychology*, *16*(3), 429–444. https://doi.org/10.1016/0193-3973(95)90028-4

McCarthy-Jones, S., Oestreich, L. K. L., Lyall, A. E., Kikinis, Z., Newell, D. T., Savadjiev, P., Shenton, M. E., Kubicki, M., Bank, A. S. R., Pasternak, O., & Whitford, T. J. (2018). Childhood adversity associated with white matter alteration in the Corpus Callosum, Corona Radiata, and Uncinate Fasciculus of psychiatrically healthy adults. *Brain Imaging and Behavior*, *12*(2), 449. https://doi.org/10.1007/s11682-017-9703-1

McKone, E., Aimola Davies, A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., & Platow, M. J. (2010). Asia has the global advantage: Race and visual attention. *Vision Research*, *50*(16), 1540–1549. https://doi.org/10.1016/j.visres.2010.05.010

McLaughlin, K. A., & Sheridan, M. A. (2016). Beyond cumulative risk: A dimensional approach to childhood adversity. *Current Directions in Psychological Science*, *25*(4), 239–245. https://doi.org/10.1177/0963721416655883

McLaughlin, K. A., Sheridan, M. A., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2021). The value of dimensional models of early experience: Thinking clearly about concepts and categories. *Perspectives on Psychological Science*, *16*(6), 1463–1472. https://doi.org/10.1177/1745691621992346

McLaughlin, K. A., Weissman, D., & Bitrán, D. (2019). Childhood adversity and neural development: A systematic review. *Annual Review of Developmental Psychology*, *1*, 277–312. https://doi.org/10.1146/annurev-devpsych-121318-084950

Mejia, A. F., Nebel, M. B., Barber, A. D., Choe, A. S., Pekar, J. J., Caffo, B. S., & Lindquist, M. A. (2018). Improved estimation of subject-level functional connectivity using full and partial correlation with empirical Bayes shrinkage. *NeuroImage*, *172*, 478–491. https://doi.org/10.1016/j.neuroimage.2018.01.029

Merkt, J., Singmann, H., Bodenburg, S., Goossens-Merkt, H., Kappes, A., Wendt, M., & Gawrilow, C. (2013). Flanker performance in female college students with ADHD: A diffusion model analysis. *ADHD Attention Deficit and Hyperactivity Disorders*, *5*(4), 321–341. https://doi.org/10.1007/s12402-013-0110-1

Merriam-Webster. (2024a). *Ability*. https://www.merriam-webster.com/dictionary/ability

Merriam-Webster. (2024b). *Performance*. https://www.merriam-webster.com/dictionary/performance

Merz, E. C., Wiltshire, C. A., & Noble, K. G. (2019). Socioeconomic inequality and the developing brain: Spotlight on language and executive function. *Child Development Perspectives*, *13*(1), 15–20. https://doi.org/10.1111/cdep.12305

Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development*, *75*(5), 1373–1386. https://doi.org/10.1111/j.1467-8624.2004.00746.x

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review*, *20*(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Miller-Cotto, D., Smith, L. V., Wang, A. H., & Ribner, A. D. (2022). Changing the conversation: A culturally responsive perspective on executive functions, minoritized children and their families. *Infant and Child Development*, *31*(1), e2286. https://doi.org/10.1002/icd.2286

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

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moos, R. H. (1994). *Family Environment Scale Manual: Development, Applications, Research*. Consulting Psychologists Press.

Muskens, M., Frankenhuis, W. E., & Borghans, L. (2019). Math items about real-world content lower test-scores of students from families with low socioeconomic status. *Npj Science of Learning*, *9*(1), 19. https://doi.org/10.1038/s41539-024-00228-8

Mutter, B., Alcorn, M., & Welsh, M. (2006). Theory of mind and executive function: Working-memory capacity and inhibitory control as predictors of false-belief task performance. *Perceptual and Motor Skills*, *102*, 819–835. https://doi.org/10.2466/PMS.102.3.819-835

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Nelson, C. A., Bhutta, Z. A., Burke Harris, N., Danese, A., & Samara, M. (2020). Adversity in childhood is linked to mental and physical health throughout life. *BMJ*, *371*, m3048. https://doi.org/10.1136/bmj.m3048

Nelson, C. A., & Gabard-Durnam, L. J. (2020). Early Adversity and Critical Periods: Neurodevelopmental Consequences of Violating the Expectable Environment. *Trends in Neurosciences*, *43*(3), 133–143. https://doi.org/10.1016/j.tins.2020.01.002

Niebaum, J. C., & Munakata, Y. (2023). Why doesn't executive function training improve academic achievement? Rethinking individual differences, relevance, and engagement from a contextual framework. *Journal of Cognition and Development*, *24*(2), 241–259. https://doi.org/10.1080/15248372.2022.2160723

Ning, K., Gondek, D., Pereira, S. M. P., & Lacey, R. E. (2023). Mediating mechanisms of the relationship between exposure to deprivation and threat during childhood and adolescent psychopathology: Evidence from the Millennium Cohort Study. *European Child & Adolescent Psychiatry*, *33*. https://doi.org/10.1007/s00787-023-02289-3

Nivison, M. D., Vandell, D. L., Booth-LaForce, C., & Roisman, G. I. (2021). Convergent and discriminant validity of retrospective assessments of the quality of childhood parenting: Prospective evidence from infancy to age 26 years. *Psychological Science*, *32*(5), 721–734. https://doi.org/10.1177/0956797620975775

Nketia, J., Al Sager, A., Dajani, R., Placido, D., & Amso, D. (2024). Executive functions in Jordanian children: What can the Hearts and Flowers Task tell us about development in a non-Western context. *Journal of Cognition and Development*, *25*(2), 180–200. https://doi.org/10.1080/15248372.2023.2248698

Noble, K. G., Hart, E. R., & Sperber, J. F. (2021). Socioeconomic disparities and neuroplasticity: Moving toward adaptation, intersectionality, and inclusion. *American Psychologist*, *76*(9), 1486–1495. https://doi.org/10.1037/amp0000934

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, *8*(1), 74–87. https://doi.org/10.1111/j.1467-7687.2005.00394.x

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*, *62*(3), 280–288. https://doi.org/10.1111/jcpp.13241

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, *134*(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

Oberauer, K. (2009). Design for a working memory. In *Psychology of Learning and Motivation* (Vol. 51, pp. 45–100). Elsevier.

Oberauer, K., & Lewandowsky, S. (2019). Simple measurement models for complex working-memory tasks. *Psychological Review*, *126*(6), 880–932. https://doi.org/10.1037/rev0000159

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. https://doi.org/10.1037/bul0000153

Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity — facets of a cognitive ability construct. *Personality and Individual Differences*, *29*(6), 1017–1045. https://doi.org/10.1016/S0191-8869(99)00251-2

Oberski, D. (2014). Lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, *57*(1), 1–27. https://doi.org/10.18637/jss.v057.i01

Parsons, S. (2021). Splithalf: Robust estimates of split half reliability. *Journal of Open Source Software*, *6*(60), 3041. https://doi.org/10.21105/joss.03041

Patten, S. B., Wilkes, T. C. R., Williams, J. V. A., Lavorato, D. H., el-Guebaly, N., Schopflocher, D., Wild, C., Colman, I., & Bulloch, A. G. M. (2015). Retrospective and prospectively assessed childhood adversity in association with major depression, alcohol consumption and painful conditions. *Epidemiology and Psychiatric Sciences*, *24*(2), 158–165. https://doi.org/10.1017/S2045796014000018

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, *51*(6), 768–774. https://doi.org/10.1002/1097-4679(199511)51:6<768::aid-jclp2270510607>3.0.co;2-1

Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational phenotyping: Using models to understand individual differences in personality, development, and mental illness. *Personality Neuroscience*, *1*, e18. https://doi.org/10.1017/pen.2018.14

Peng, P., & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain? *Journal of Learning Disabilities*, *49*(1), 3–20. https://doi.org/10.1177/0022219414521667

Pirrone, A., Stafford, T., & Marshall, J. A. R. (2014). When natural selection should optimize speed-accuracy trade-offs. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00073

Polavarapu, A., & Hasbani, D. (2017). Neurological complications of nutritional disease. *Seminars in Pediatric Neurology*, *24*(1), 70–80. https://doi.org/10.1016/j.spen.2016.12.002

Pollak, S. D. (2008). Mechanisms linking early experience and the emergence of emotions: Illustrations from the study of maltreated children. *Current Directions in Psychological Science*, *17*(6), 370–375. https://doi.org/10.1111/j.1467-8721.2008.00608.x

Poole, D., Grange, J. A., & Milne, E. (2024). Putting the spotlight back onto the flanker task in autism: Autistic adults show increased interference from foils compared with non-autistic adults. *Journal of Cognition*, *7*(1), 1–15. https://doi.org/10.5334/joc.369

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(2), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychological Review*, *74*(5), 392–409. https://doi.org/10.1037/h0024913

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision*, *2*(4), 237–279. https://doi.org/10.1037/dec0000030

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., Smith, P. L., & McKoon, G. (2015). Modeling regularities in response time and accuracy data with the Diffusion Model. *Current Directions in Psychological Science*, *24*(6), 458–470. https://doi.org/10.1177/0963721415596228

Rebello, K., Moura, L. M., Pinaya, W. H. L., Rohde, L. A., & Sato, J. R. (2018). Default mode network maturation and environmental adversities during childhood. *Chronic Stress*, *2*, 2470547018808295. https://doi.org/10.1177/2470547018808295

Rey-Mermet, A. (2024). *Beyond unity and diversity: Tasks used to assess attentional control measure task-specific processes*. OSF. https://doi.org/10.31234/osf.io/jvbfn

Rey-Mermet, A., Gade, M., Souza, A. S., Bastian, C. C. von, & Oberauer, K. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology. General*, *148*(8), 1335–1372. https://doi.org/10.1037/xge0000593

Reynolds, A. J., Ou, S.-R., Mondi, C. F., & Giovanelli, A. (2019). Reducing poverty and inequality through preschool-to-3rd grade prevention services. *The American Psychologist*, *74*(6), 653. https://doi.org/10.1037/amp0000537

Ribeiro, M., Yordanova, Y. N., Noblet, V., Herbet, G., & Ricard, D. (2024). White matter tracts and executive functions: A review of causal and correlation evidence. *Brain*, *147*(2), 352–371. https://doi.org/10.1093/brain/awad308

Ridderinkhof, K. R., Wylie, S. A., Wildenberg, W. P. M. van den, Bashore, T. R., & Molen, M. W. van der. (2021). The arrow of time: Advancing insights into action control from the arrow version of the Eriksen flanker task. *Attention, Perception, & Psychophysics*, *83*(2), 700–721. https://doi.org/10.3758/s13414-020-02167-z

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*, *12*(6), 890–901. https://doi.org/10.1017/S2040174420001348

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rosen, M. L., Amso, D., & McLaughlin, K. A. (2019). The role of the visual association cortex in scaffolding prefrontal cortex development: A novel mechanism linking socioeconomic status and executive function. *Developmental Cognitive Neuroscience*, *39*, 100699. https://doi.org/10.1016/j.dcn.2019.100699

Rosen, M. L., Sheridan, M. A., Sambrook, K. A., Meltzoff, A. N., & McLaughlin, K. A. (2018). Socioeconomic disparities in academic achievement: A multi-modal investigation of neural mechanisms in children and adolescents. *NeuroImage*, *173*, 298–310. https://doi.org/10.1016/j.neuroimage.2018.02.043

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Salhi, C., Beatriz, E., McBain, R., McCoy, D., Sheridan, M., & Fink, G. (2021). Physical discipline, deprivation, and differential risk of developmental delay across 17 countries.

*Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(2), 296–306. https://doi.org/10.1016/j.jaac.2020.02.016

Salthouse, T. A. (2016). Little relation of adult age on cognition after controlling general influences. *Developmental Psychology*, *52*(10), 1545–1554. https://doi.org/10.1037/dev0000162

Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, *34*(1), 17–24. https://doi.org/10.1037/pag0000288

Schäfer, J. L., McLaughlin, K. A., Manfro, G. G., Pan, P., Rohde, L. A., Miguel, E. C., Simioni, A., Hoffmann, M. S., & Salum, G. A. (2022). Threat and deprivation are associated with distinct aspects of cognition, emotional processing, and psychopathology in children and adolescents. *Developmental Science*, *26*(1), e13267. https://doi.org/10.1111/desc.13267

Scherpenzeel, A. (2011). Data collection in a probability-based internet panel: How the LISS Panel was built and how it can be used. *Bulletin of Sociological Methodology*, *109*, 56–61. https://doi.org/10.1177/0759106310387713

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(4), 1089–1096. https://doi.org/10.1037/a0015730

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414–429. https://doi.org/10.1037/0096-3445.136.3.414

Schmitz, F., & Wilhelm, O. (2016). Modeling mental speed: Decomposing response time distributions in elementary cognitive tasks and correlations with working memory capacity and fluid intelligence. *Journal of Intelligence*, *4*(4), 13. https://doi.org/10.3390/jintelligence4040013

Schreiter, S., & Vogel, M. (2023). Eye-tracking measures as indicators for a local vs. Global view of data. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.1058150

Schreiter, S., & Vogel, M. (2024). Students' local vs. Global views of data distributions: A cross-grade-level analysis using eye-tracking. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-024-10352-2

Schubert, A.-L., & Frischkorn, G. T. (2020). Neurocognitive psychometrics of intelligence: How measurement advancements unveiled the role of mental speed in intelligence differences. *Current Directions in Psychological Science*, *29*(2), 140–146. https://doi.org/10.1177/0963721419896365

Schubert, A.-L., Frischkorn, G. T., Hagemann, D., & Voss, A. (2016). Trait characteristics of diffusion model parameters. *Journal of Intelligence*, *4*(3), 7. https://doi.org/10.3390/jintelligence4030007

Schubert, A.-L., Nunez, M. D., Hagemann, D., & Vandekerckhove, J. (2019). Individual differences in cortical processing speed predict cognitive abilities: A model-based cognitive neuroscience account. *Computational Brain & Behavior*, *2*(2), 64–84. https://doi.org/10.1007/s42113-018-0021-5

Schwabe, L., Höffken, O., Tegenthoff, M., & Wolf, O. T. (2013). Stress-induced enhancement of response inhibition depends on mineralocorticoid receptor activation. *Psychoneuroendocrinology*, *38*(10), 2319–2326. https://doi.org/10.1016/j.psyneuen.2013.05.001

Schwabe, L., Oitzl, M. S., Philippsen, C., Richter, S., Bohringer, A., Wippich, W., & Schachinger, H. (2007). Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Learning & Memory*, *14*(1-2), 109–116. https://doi.org/10.1101/lm.435807

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience*, *29*(22), 7191–7198. https://www.jneurosci.org/content/29/22/7191.short

Shariq, D., Romeo, R. R., & Gard, A. M. (2024). *Cognitive resilience and vulnerability to socioeconomic disadvantage: Individual, family, school, and neighborhood predictors*. OSF. https://doi.org/10.31234/osf.io/2kew7

Sheridan, M. A., & McLaughlin, K. A. (2014). Dimensions of early experience and neural development: Deprivation and threat. *Trends in Cognitive Sciences*, *18*(11), 580–585. https://doi.org/10.1016/j.tics.2014.09.001

Sheridan, M. A., Mukerji, C. E., Wade, M., Humphreys, K. L., Garrisi, K., Goel, S., Patel, K., Fox, N. A., Zeanah, C. H., Nelson, C. A., & McLaughlin, K. A. (2022). Early deprivation alters structural brain development from middle childhood to adolescence. *Science Advances*, *8*(40), eabn4316. https://doi.org/10.1126/sciadv.abn4316

Sheridan, M. A., Shi, F., Miller, A. B., Salhi, C., & McLaughlin, K. A. (2020). Network structure reveals clusters of associations between childhood adversities and development outcomes. *Developmental Science*, *23*(5), e12934. https://doi.org/10.1111/desc.12934

Shields, G. S., Bonner, J. C., & Moons, W. G. (2015). Does cortisol influence core executive functions? A meta-analysis of acute cortisol administration effects on working memory, inhibition, and set-shifting. *Psychoneuroendocrinology*, *58*, 91–103. https://doi.org/10.1016/j.psyneuen.2015.04.017

Shields, G. S., Doty, D., Shields, R. H., Gower, G., Slavich, G. M., & Yonelinas, A. P. (2017). Recent life stress exposure is associated with poorer long-term memory, working memory, and self-reported memory. *Stress*, *20*(6), 598–607. https://doi.org/10.1080/10253890.2017.1380620

Shields, G. S., Ramey, M. M., Slavich, G. M., & Yonelinas, A. P. (2019). Determining the mechanisms through which recent life stress predicts working memory impairments: Precision or capacity? *Stress*, *22*(2), 280–285. https://doi.org/10.1080/10253890.2018.1556635

Shields, G. S., Sazma, M. A., & Yonelinas, A. P. (2016). The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience and Biobehavioral Reviews*, *68*, 651–668. https://doi.org/10.1016/j.neubiorev.2016.06.038

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*(6), 771–799. https://doi.org/10.1177/1745691616650647

Shonkoff, J. P., Garner, A. S., Child, C. on P. A. of, Health, F., Care, D., Developmental, S. on, Pediatrics, B., Siegel, B. S., Dobbins, M. I., Earls, M. F., Garner, A. S., McGuinn, L., Pascoe, J., & Wood, D. L. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*, *129*(1), e232–e246. https://doi.org/10.1542/peds.2011-2663

Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics*, *6*(1), 99–105. https://doi.org/10.1080/00140136308930679

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. *Nature Human Behaviour*, *4*(11), 1208–1214. https://doi.org/10.1038/s41562-020-0912-z

Sliwinski, M. J., Smyth, J. M., Hofer, S. M., & Stawski, R. S. (2006). Intraindividual coupling of daily stress and cognition. *Psychology and Aging*, *21*(3), 545–557. https://doi.org/10.1037/0882-7974.21.3.545

Slotkin, J., Kallen, M., Griffith, J., Magasi, S., Salsman, J., & Nowinski, C. (2012). *NIH toolbox: Scoring and interpretation guide*.

Smith, K. E., & Pollak, S. D. (2021). Rethinking concepts and categories for understanding the neurodevelopmental effects of childhood adversity. *Perspectives on Psychological Science*, *16*(1), 67–93. https://doi.org/10.1177/1745691620920725

Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychological Bulletin*, *147*(4), 329–351. https://doi.org/10.1037/bul0000322

Spielberger, C. D., Sydeman, S. J., Owen, A. E., & Marsh, B. J. (1999). Measuring anxiety and anger with the State-Trait Anxiety Inventory (STAI) and the State-Trait Anger Expression Inventory (STAXI). In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (pp. 993–1021). Lawrence Erlbaum Associates Publishers.

Spinella, M. (2007). Normative data and a short form of the Barratt Impulsiveness Scale. *International Journal of Neuroscience*, *117*(3), 359–368. https://doi.org/10.1080/00207450600588881

Stahl, C., Voss, A., Schmitz, F., Nuszbaum, M., Tüscher, O., Lieb, K., & Klauer, K. C. (2014). Behavioral components of impulsivity. *Journal of Experimental Psychology. General*, *143*(2), 850–886. https://doi.org/10.1037/a0033981

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed–accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*(2), 377–390. https://doi.org/10.1037/a0018022

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Development*, *80*(1), 28–44. https://doi.org/10.1111/j.1467-8624.2008.01244.x

Steudte-Schmiedgen, S., Stalder, T., Kirschbaum, C., Weber, F., Hoyer, J., & Plessow, F. (2014). Trauma exposure is associated with increased context-dependent adjustments of cognitive control in patients with posttraumatic stress disorder and healthy controls. *Cognitive, Affective & Behavioral Neuroscience*, *14*(4), 1310–1319. https://doi.org/10.3758/s13415-014-0299-2

Stevens, C., Lauinger, B., & Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: An event-related brain potential study. *Developmental Science*, *12*(4), 634–646. https://doi.org/10.1111/j.1467-7687.2009.00807.x

Sweeten, G. (2012). Scaling criminal offending. *Journal of Quantitative Criminology*, *28*(3), 533–557. https://doi.org/10.1007/s10940-011-9160-8

Taylor, R. L., & Barch, D. M. (2022). Inhibitory control within the context of early life poverty and implications for outcomes. *Neuroscience & Biobehavioral Reviews*, *140*, 104778. https://doi.org/10.1016/j.neubiorev.2022.104778

Team, R. C. (2022). *R: A language and environment for statistical computing.* https://www.R-project.org/

Textor, J., Zander, B. van der, Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: The R package "dagitty." *International Journal of Epidemiology*, *45*(6), 1887–1894. https://doi.org/10.1093/ije/dyw341

Thompson, A., & Steinbeis, N. (2021). Computational modelling of attentional bias towards threat in paediatric anxiety. *Developmental Science*, *24*(3), e13055. https://doi.org/10.1111/desc.13055

Tibu, F., Sheridan, M. A., McLaughlin, K. A., Nelson, C. A., Fox, N. A., & Zeanah, C. H. (2016). Disruptions of working memory and inhibition mediate the association between exposure to institutionalization and symptoms of attention deficit hyperactivity disorder. *Psychological Medicine*, *46*(3), 529–541. https://doi.org/10.1017/S0033291715002020

Trimmer, P. C., Houston, A. I., Marshall, J. A. R., Bogacz, R., Paul, E. S., Mendl, M. T., & McNamara, J. M. (2008). Mammalian choices: Combining fast-but-inaccurate and slow-but-accurate decision-making systems. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1649), 2353–2361. https://doi.org/10.1098/rspb.2008.0417

Tucker-Drob, E. M. (2013). How many pathways underlie socioeconomic differences in the development of cognition and achievement? *Learning and Individual Differences*, *25*, 12–20. https://doi.org/10.1016/j.lindif.2013.01.015

Ugarte, E., & Hastings, P. D. (2023). Assessing unpredictability in caregiver–child relationships: Insights from theoretical and empirical perspectives. *Development and Psychopathology*, *36*(3), 1070–1089. https://doi.org/10.1017/S0954579423000305

Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2015). Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology*, *78*, 148–174. https://doi.org/10.1016/j.cogpsych.2015.02.005

Ursache, A., & Noble, K. G. (2016a). Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology*, *53*(1), 71–82. https://doi.org/10.1111/psyp.12547

Ursache, A., & Noble, K. G. (2016b). Socioeconomic status, white matter, and executive function in children. *Brain and Behavior*, *6*(10), e00531. https://doi.org/10.1002/brb3.531

Van Calster, L., D'Argembeau, A., & Majerus, S. (2018). Measuring individual differences in internal versus external attention: The attentional style questionnaire. *Personality and Individual Differences*, *128*, 25–32. https://doi.org/10.1016/j.paid.2018.02.014

Van den Brakel, M., Lok, R., Otten, F., Vandewal, E., Bos, J., Warnaar, M., Wieman, G., Goderis, B., Hoff, S., Muns, S., & Tunderman, S. (2023). *Op weg naar een nieuwe armoedegrens. Tussenrapport van het gezamenlijke project 'Uniformering armoedeafbakening'.* https://www.scp.nl/publicaties/publicaties/2023/06/30/op-weg-naar-een-nieuwe-armoedegrens

Van Gelder, J.-L., & Frankenhuis, W. E. (2024). Short-term mindsets and crime. *Annual Review of Criminology*, *8*, 15.1–15.26. https://doi.org/10.1146/annurev-criminol-022422-124536

Van Veen, V., Krug, M. K., & Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *Journal of Cognitive Neuroscience*, *20*(11), 1952–1965. https://doi.org/10.1162/jocn.2008.20146

Vandekerckhove, J. (2014). A cognitive latent variable model for the simultaneous analysis of behavioral and personality data. *Journal of Mathematical Psychology*, *60*, 58–71. https://doi.org/10.1016/j.jmp.2014.06.004

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44–62. https://doi.org/10.1037/a0021765

Vermeent, S. (2023). *Projectlog: Tools for documenting your project workflow*. https://stefanvermeent.github.io/projectlog/

Vermeent, S., Schubert, A.-L., DeJoseph, M. L., Denissen, J. J. A., Van Gelder, J.-L., & Frankenhuis, W. E. (2025). Inconclusive evidence for associations between adverse experiences in adulthood and working memory performance. *Royal Society Open Science*. https://doi.org/10.1098/rsos.241837

Vermeent, S., Schubert, A.-L., Young, E. S., DeJoseph, M. L., & Frankenhuis, W. E. (2024). Cognitive deficits and enhancements in youth from adverse conditions: An integrative assessment using Drift Diffusion Modeling in the ABCD study. *Developmental Science*, *e13478*. https://doi.org/10.1111/desc.13478

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental Psychology*, *60*(6), 385–402. https://doi.org/10.1027/1618-3169/a000218

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/BF03196893

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00336

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*(5), 641–671. https://doi.org/10.1080/09541440802205067

Walasek, N., Young, E. S., & Frankenhuis, W. E. (2024). A framework for studying environmental statistics in developmental science. *Psychological Methods*. https://doi.org/10.1037/met0000651

Weigard, A., Clark, D. A., & Sripada, C. (2021). Cognitive efficiency beats top-down control as a reliable individual difference dimension relevant to self-control. *Cognition*, *215*, 104818. https://doi.org/10.1016/j.cognition.2021.104818

Weigard, A., & Sripada, C. (2021). Task-general efficiency of evidence accumulation as a computationally defined neurocognitive trait: Implications for clinical neuroscience. *Biological Psychiatry Global Open Science*, *1*(1), 5–15. https://doi.org/10.1016/j.bpsgos.2021.02.001

White, C. N., & Curl, R. (2018). Cueing effects in the Attentional Network Test: A Spotlight Diffusion Model analysis. *Computational Brain & Behavior*, *1*(1), 59–68. https://doi.org/10.1007/s42113-018-0004-6

White, C. N., & Kitchen, K. N. (2022). On the need to improve the way individual differences in cognitive function are measured with reaction time tasks. *Current Directions in Psychological Science*, *31*(3), 223–230. https://doi.org/10.1177/09637214221077060

White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive Psychology*, *63*(4), 210–238. https://doi.org/10.1016/j.cogpsych.2011.08.001

White, C. N., Servant, M., & Logan, G. D. (2018). Testing the validity of conflict drift-diffusion models for use in estimating cognitive processes: A parameter-recovery study. *Psychonomic Bulletin & Review*, *25*(1), 286–301. https://doi.org/10.3758/s13423-017-1271-2

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*, 14. https://doi.org/10.3389/fninf.2013.00014

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00433

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*, 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. (2018). Can an unpredictable childhood environment enhance working memory? Testing the sensitizedspecialization hypothesis. *Journal of Personality and Social Psychology*, *114*(6), 891–908. https://doi.org/10.1037/pspi0000124

Young, E. S., & Vermeent, S. (2023). *Multitool: Run multiverse style analyses*. https://ethan-young.github.io/multitool/

Young, E. S., Vermeent, S., Frankenhuis, W. E., Nivison, M., Simpson, J. A., & Roisman, G. I. (2024). How does adversity shape performance across different abilities in the same person? *Development and Psychopathology*, 1–18. https://doi.org/10.1017/S0954579424001433

Yuan, Z., Qin, W., Wang, D., Jiang, T., Zhang, Y., & Yu, C. (2012). The salience network contributes to an individual's fluid reasoning capacity. *Behavioural Brain Research*, *229*(2), 384–390. https://doi.org/10.1016/j.bbr.2012.01.037

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*(1), 297–301. https://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., Conway, K. P., Gershon, R., & Weintraub, S. (2014). NIH Toolbox Cognition Battery (CB): Validation of executive function measures in adults. *Journal of the International Neuropsychological Society*, *20*(6), 620–629. https://doi.org/10.1017/S1355617714000472

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. https://doi.org/10.1111/mono.12032

Zelazo, P. D., & Carlson, S. M. (2023). Reconciling the context-dependency and domain-generality of executive function skills from a developmental systems perspective. *Journal of Cognition and Development*, *24*(2), 205–222. https://doi.org/10.1080/15248372.2022.2156515

Zucker, R. A., Gonzalez, R., Feldstein Ewing, S. W., Paulus, M. P., Arroyo, J., Fuligni, A., Morris, A. S., Sanchez, M., & Wills, T. (2018). Assessment of culture and environment in the Adolescent Brain and Cognitive Development Study: Rationale, description of measures, and early data. *Developmental Cognitive Neuroscience*, *32*, 107–120. https://doi.org/10.1016/j.dcn.2018.03.004

Zuilkowski, S. S., McCoy, D. C., Serpell, R., Matafwali, B., & Fink, G. (2016). Dimensionality and the development of cognitive assessments for children in Sub-Saharan

Africa. *Journal of Cross-Cultural Psychology*, *47*(3), 341–354. https://doi.org/10.1177/0022022115624155

## Supplementary materials

# Appendix 1 - Chapter 2

## Data access workflow

Prior to Stage 1 submission of the Registered Report, we accessed the cognitive task data for a couple of preregistered data checks. By only accessing the cognitive task data, these steps did not bias or substantive analyses involving measures of adversity. To transparently show when we accessed which data, we created an open science workflow that would automate this process. The main aim of this workflow was to create a transparent log of every major milestone of the project, such as accessing new data, submitting preregistrations, and finalizing analyses.

The main ingredient of this workflow is a set of custom functions that we created for reading in data files (See Figure A1.1). These are wrappers for the read functions in the *readr* package. Whenever one of these functions (e.g., *read_csv*) was called, it went through a couple of internal processes. First, the specified data file would be read into R (but not yet accessible to us in the global environment). This could be a single file, or a list of individual data files that would first be combined into a single dataframe. Second, any specified manipulations would be applied to the data. This could be selecting specific variables, filtering specific rows, or randomly shuffling values (e.g., participant IDs). Third, An MD5 hash of the final R object would be generated using the *digest* package. An MD5 hash is a unique, 32-digit string that maps directly onto the content of the R object. The same R object will always generate the same MD5 hash, but as soon as anything changes (e.g., a variable is added, a value is rounded), the MD5 hash changes. Fourth, this MD5 hash would be compared to previously generated hashes.

If the newly generated MD5 hash was not recognized, this triggered an automatic commit to GitHub. At this point, the user gets the choice to abort the process or to continue. Aborting would terminate the process without importing the data. If opting to continue, the user could supply an informative message (e.g., “accessed Flanker data”), which would be added to the Git commit. The Git commit message stored other relevant meta-data as well, such as the object hash and the code used to read and manipulate the data. Committing and pushing to Git was handled using the *gert* package.

Thus, any accessing of raw data was automatically tracked via GitHub. Using this same approach, we also logged other major milestones, such as submitting preregistrations and finalizing analyses.

An automatically generated overview of all milestones can be found in the Data Access History.

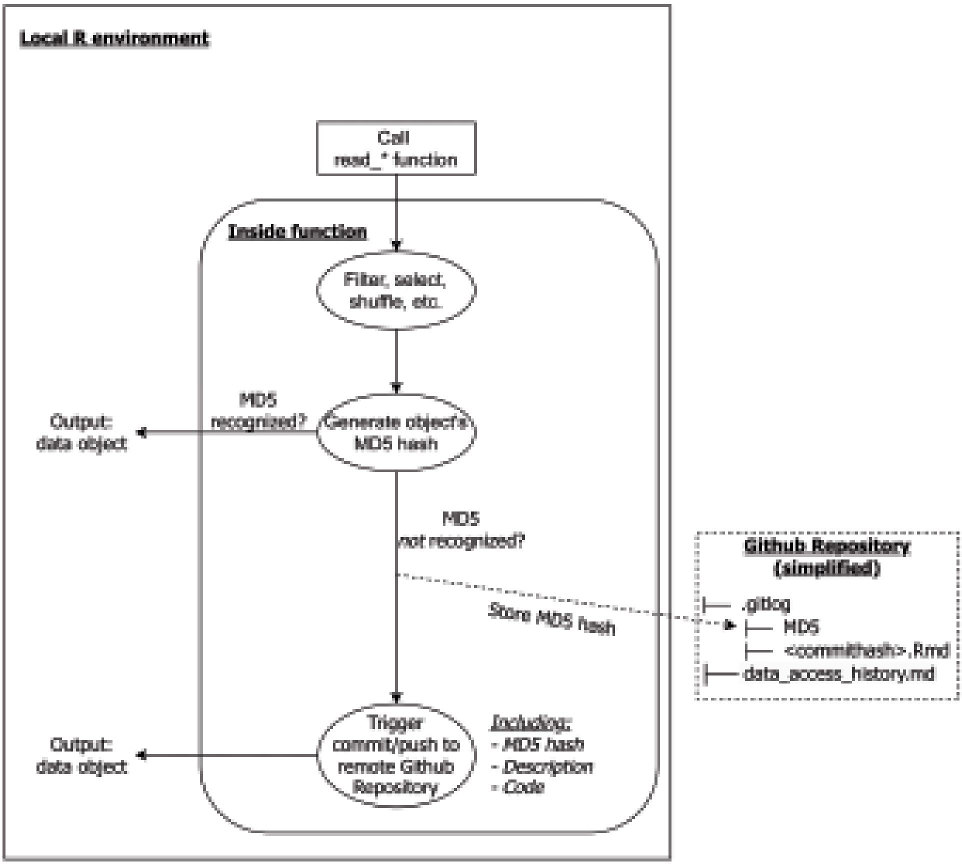



**Figure A1.1.** Graphical overview of the data access workflow using R and GitHub.

## Power analysis

We conducted a power analysis through simulation using the *simulateData* function of the *lavaan* package. On each iteration, we first specified a population model (i.e., the 'true' model) with prespecified factor loadings and regression coefficients. Factor loadings in this model were randomly generated between 0.6 and 0.8 following a uniform distribution. Next, we simulated data sets based on the population model. Finally, we fitted a sample model (i.e., without constrained parameters) to the simulated data and extracted the beta coefficients and corresponding *p*-values. We generated population models with beta coefficients of 0.06, 0.08 and 0.1, and simulated data with sample sizes ranging from 1,500 to 8,500 with steps of 1,000. Each combination of coefficients and sample sizes was repeated 500 times, for a total of 12,000 iterations.

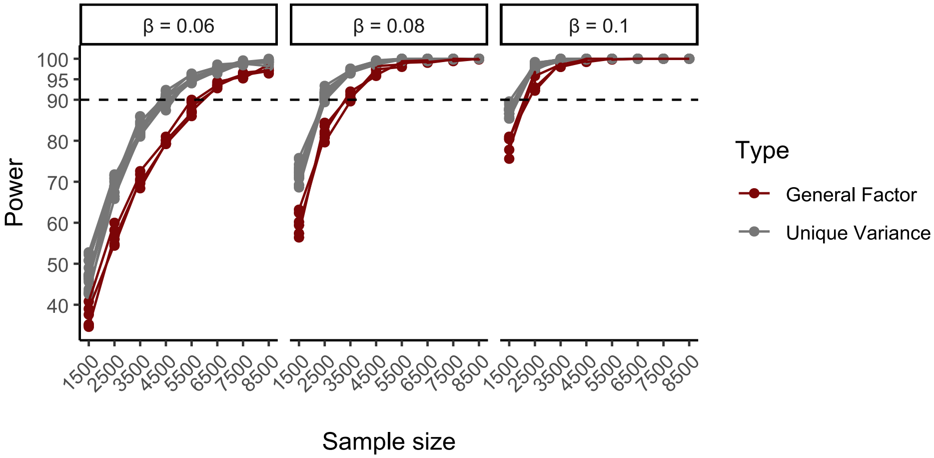



**Figure A1.2.** Results of the power simulations. The dashed line indicates 90% power.

The results are shown in Figure A1.2. The simulations yield power > 90% at around $N$ = 2,500 for $\beta$ = 0.1 and $N$ = 6,500 for $\beta$ = 0.06. Thus, after taking out 1,500 participants for the training set, the test set is highly powered.

## Response Distributions of Cognitive Tasks

See Table A1.1 for descriptive statistics for all cognitive tasks.

**Table A1.1.**
Descriptive statistics of mean RTs and accuracy for the cognitive tasks.

| | RT Mean (SD) | Accuracy Mean (SD) | Accuracy Min | Accuracy Max |
|---|---|---|---|---|
| Processing Speed | 2.24 (0.47) | 96.42 (4.3) | 55.17 | 100 |
| Flanker | 0.91 (0.33) | 99.31 (3.25) | 52.63 | 100 |
| Mental Rotation | 2.65 (0.47) | 59.25 (16.81) | 6.25 | 100 |
| Attention Shifting | 1.01 (0.35) | 92.94 (6.76) | 22.22 | 100 |

## Overview of DDM Modeling Procedure

In theory, the hierarchical Bayesian framework allows simultaneously estimating DDM parameters, latent measurement models, and the regression paths between them in a single step (e.g., Schubert et al., 2019; Vandekerckhove, 2014). An advantage of this approach is that information regarding estimation uncertainty (e.g., of the DDM parameters) gets integrated in subsequent steps. However, this approach is very computationally expensive and might even be unfeasible with the current sample size. Therefore, we opted for a two-step estimation approach.

The hierarchical DDM models was fit using the *runjags* package (Denwood, 2016) with JAGS code adapted from D. J. Johnson et al. (2017). The JAGS code was adjusted in a number of ways to meet our purposes. Across all models, the starting point was fixed to 0.5, and the boundary separation was constrained to be the same across conditions where relevant. Each model was fit with three Markov Chain Monte Carlo (M-CMC) chains. Each chain contained 2,000 burn-in samples and 10,000 additional samples. Of these samples, every 10th sample was retained. Posterior samples of all three chains were combined, resulting in a posterior sample of 3,000 samples. If a model did not converge properly with these settings, we increased the amount of samples drawn stepwise up to 100,000.

Model convergence was assessed in several ways. First, we visually inspected the traces, which should not contain any drifts or large jumps. Second, we calculated the Gelman-Rubin convergence statistic R^ (Gelman & Rubin, 1992), of which all values should be below 1.1. Third, we assessed whether the model provided a good fit to the participants' data using simulation (See Figure A1.3 for a visualization of this procedure). When we estimate DDM parameters for a participant, we want to be sufficiently sure that the parameters accurately reflect the participant's real cognitive processes. Some factors can bias estimates. For example, trial-level outliers could bias DDM parameters so that they are no longer representative of the full RT distribution. Thus, before using the obtained DDM parameters to address our hypotheses, we need to make sure that they accurately reflect participant's cognitive processes. It is standard practice in cognitive modeling to use simulation to evaluate the accuracy of parameter recovery (Lewandowsky & Farrell, 2010). Imagine that for child A, the model estimates a drift rate of 2, a boundary separation of 1, and a non-decision time of 0.5. To evaluate whether these values likely reflect the child's "true" parameter values (i.e., the combination of cognitive processes that produce their pattern of RTs and accuracy), we take each child's estimated DDM parameters and use them to simulate RT/accuracy data. This procedure is analogous to drawing values from a normal distribution if we know the relevant parameters (i.e., the mean and standard deviation). Similarly, we can draw simulated values (combinations of an RT and accuracy) based on the child's parameter estimates. If the child's DDM parameter estimates are valid, the simulated RT/accuracy data should be highly correlated with the child's actual data. We computed overall correlations between the observed and simulated scores for RTs in the 25th, 50th and 75th percentile of the RT distribution as well as for accuracy rates. If the correlation was < .80, we took steps to improve model fit (see below).

In addition, we also computed correlations between observed and simulated RTs and accuracy at different levels of the two adversity measures: <1*SD*, ≥1*SD*≤, and >1*SD*. This told us whether parameter recovery was worse for specific subgroups of participants, which would require caution when interpreting the results. If correlations for specific subgroups were low but the overall correlation was > .80, we still used the estimates in the analyses.

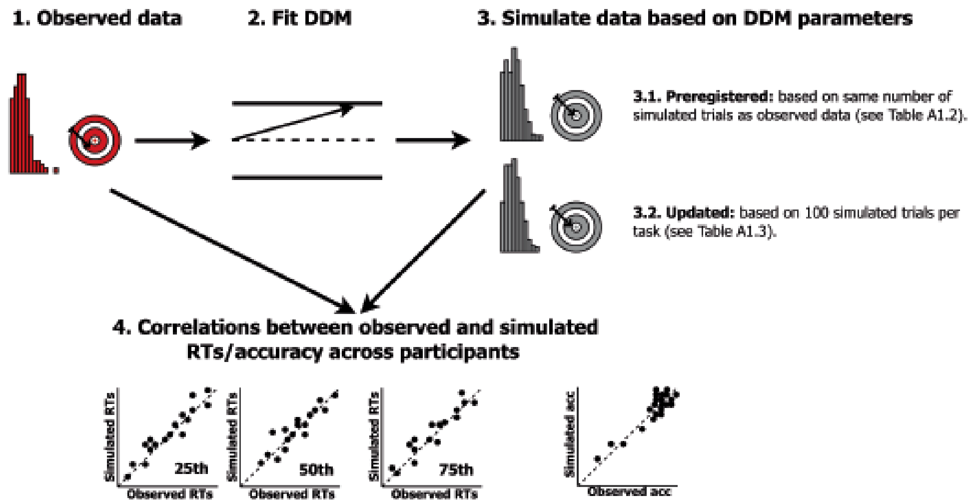



**Figure A1.3.** Graphical overview of the simulation-based model fit procedure. First, we fit the DDM to the observed response times and accuracy rates (step 1-2). Then, we use the resulting DDM parameter estimates of each participant to simulate new data (step 3). Finally, we compute correlations between the observed and simulated response times (separately at the 25th, 50h, and 75th percentile of the response time distribution) and accuracy rates (Step 4). We deviated from our preregistered simulation procedure (simulating the same number of trials as the observed data; step 3.1) by instead simulating 100 trials per task (step 3.2). This deviation is explained in more detail in the main text. Note: The scatterplots do not present real data but are for illustrative purposes only.

In case of overall model fit < .80 for a particular task, we determined criteria to find outliers based on the following simulation procedure. First, we would simulate DDM parameters for 10,000 participants based on the overall sample parameter distributions (means, standard deviations, and the variance-covariance matrix). Second, we would generate RT and accuracy data based on this new set of simulated parameters. Third, we would fit the DDM to these RT and accuracy data and again generate RT and accuracy data from these estimated DDM parameters. Thus, this procedure would yield a set of simulated RT/accuracy data and corresponding recovered RT/accuracy data. We would fit regression models predicting estimated RTs and accuracy with simulated RTs and accuracy at the 25th, 50th and 75th percentile. The 2.5% and 97.5% quantiles of the residuals would be extracted from each model and used as cut-offs for bad model fit. Participants would be excluded if any of their RTs or accuracies are larger than these cut-offs. After excluding outliers, we would fit the DDM model again and repeat model fit assessments.

### Imputation of the Mental Rotation Task

During preprocessing, we discovered that the 5-second response cut-off that was used for the Mental Rotation Task led to severe truncation of the RT distribution. This is problematic because the tail of the distribution holds important information about stages of processing. Truncation of reasonably long RTs can therefore lead to biased DDM parameter estimates. The hierarchical Bayesian framework allows these missing values to be imputed based on the rest of the data, which has been shown to lead to unbiased estimates. The procedure is described in detail in the supplemental materials of D. J. Johnson et al. (2017). In short, it involves two steps. First, responses are sampled probabilistically for each missing trial based on the overall accuracy of the participant. For example, if a participant has an overall accuracy of 80%, each missing response has a probability of .80 to be assigned a 1 (i.e., correct response). Second, responses are assigned to three bins. The first bin contains incorrect (imputed) RTs slower than 5 seconds (coded as −5). The second bin contains the observed data, ranging between −5 and 5 seconds. The third bin contains correct (imputed) RTs slower than 5 seconds (coded as 5). JAGS then imputes the response times for missing trials based on these thresholds. We will compare model versions with and without imputation of missing responses. A simulation demonstrating the feasibility of this approach is described below (DDM simulation 5: Imputation of missing RTs)

### DDM simulations: The effect of few trials per participant

The number of trials that is available for each of the cognitive tasks is substantially lower than is typical for DDM analyses. This is especially true for the Flanker Task (8 incongruent trials, 12 congruent trials) and the Attention Shifting Task (7 switch trials, 23 repeat trials). While each participant completed a small number of trials, the hierarchical Bayesian framework can use information from the full sample to estimate and constrain individual estimates. Here, we report simulation studies that aimed to assess whether it would be possible to accurately recover parameter estimates. The analyses are modeled on the Flanker Task, which is the task with the lowest overall number of trials ($N$ = 20). For simulations involving two conditions, we assume (as we do in the real data) that the drift rate and non-decision time differ (and are correlated) across conditions, and that the boundary separation is the same across conditions. This latter assumption reflects the fact that conditions are randomly shuffled on a trial-by-trial basis, which prohibits participants from adapting their strategy for different conditions. The starting point is fixed to the mid-point (0.5) for all simulations.

#### DDM simulation 1: Single condition with eight trials

First, we simulated task data for 1,500 participants with eight trials per participant. We used the first 2,000 samples as burn-in, and then took an additional 10,000 samples. Every 10th sample was discarded to reduce the size of final model object. We sampled across three chains, which were subsequently combined, for a total of 3,000 samples. The model converged normally (Figure A1.4). Relative parameter recovery was decent for boundary separation ($r$ = .76) and non-decision time ($r$ = .73), but not for drift rate ($r$ = .54). However, estimates of boundary separation and non-decision time showed substantial bias (See Figure A1.5).

As discussed above, the models reported in this manuscript will not be constricted to eight trials. Instead, they will be able to use the information of both conditions (e.g.,

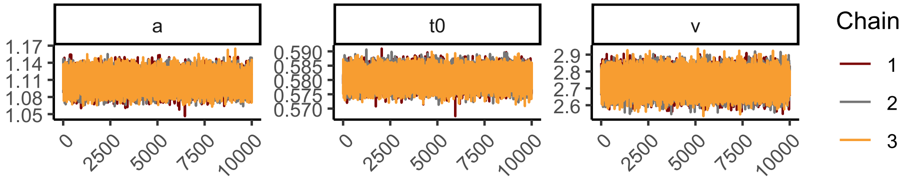



**Figure A1.4.** Convergence of the model in simulation 1. Plots should resemble a 'fat, hairy caterpillar'.

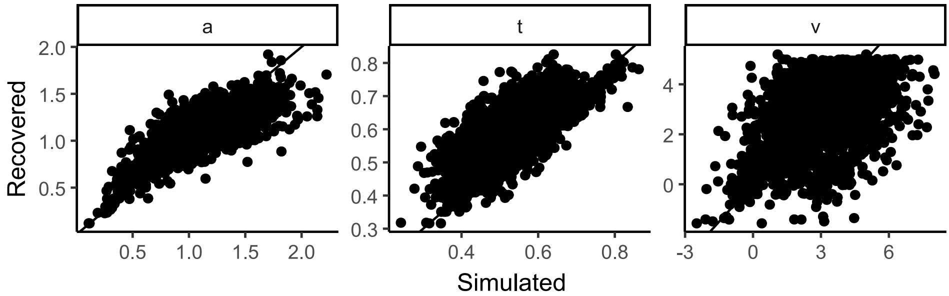



**Figure A1.5.** Parameter recovery in the case of two conditions. a = Boundary Separation; t = Non-Decision Time; v = Drift Rate.

congruent and incongruent for the Flanker Task), as parameters will tend to be correlated across conditions. Therefore, we ran a second simulation adding realistic condition effects.

**DDM simulation 2: Two conditions; Boundary Separation fixed across conditions**

We again simulated task data for 1,500 participants. Mirroring the real Flanker task, we simulated two conditions, one with 8 trials (incongruent) and one with 12 trials (congruent). On average, drift rates were lower and non-decision times were longer for incongruent trials. Boundary separation was fixed within subjects to be equal across conditions. Non-decision times correlated on average .70 between conditions, and drift rates correlated on average .30 between conditions. These correlations were based on previous studies that we did using the Flanker Task. For more information on the specific settings, see https://github.com/stefanvermeent/abcd_ddm/scripts/0_simulations/ddm_trial_simulations.R.

As the model converged without issues in simulation 1, we tried reducing the number of samples (2,000 burn-in with an additional 2,000 samples) to save time. The model converged normally (Figure A1.6). Correlations between simulated and recov-

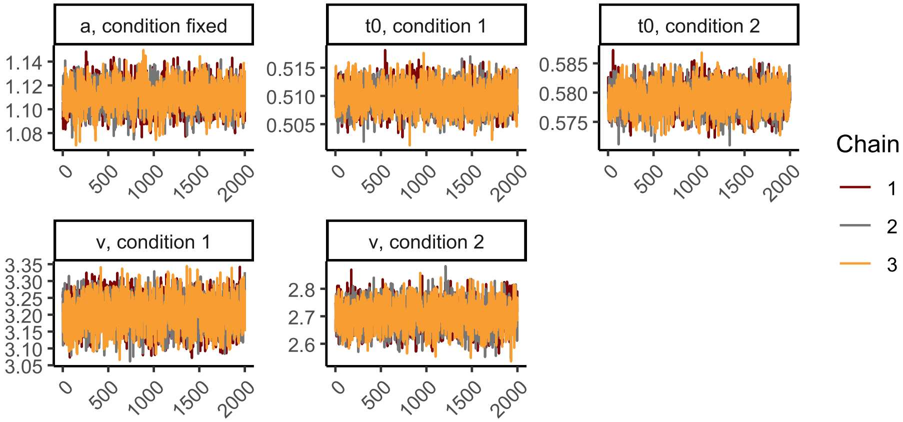



**Figure A1.6.** Convergence of the model in simulation 2. Plots should resemble a 'fat, hairy caterpillar'.

ered parameter estimates was high, ranging between $r$ = .84 for the drift rate and .95 for the non-decision time (see Figure A1.7).

Simulation 1 and 2 involved data of 1,500 simulated subjects. However, the sample size of our real data set is roughly 10,000. Thus, in the real data there is substantially more group-level data to inform and constrain the individual parameter estimates. we ran a third simulation to investigate if—and to what extent—the parameter estimates would improve moving from 1,500 to 10,000 participants.

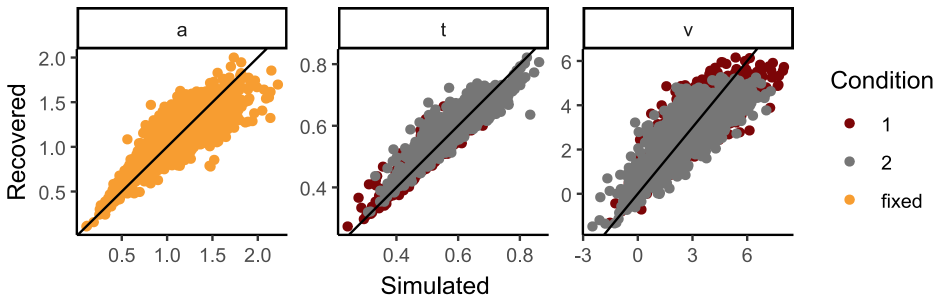



**Figure A1.7.** Parameter recovery in the case of two conditions. a = Boundary Separation; t = Non-Decision Time; v = Drift Rate

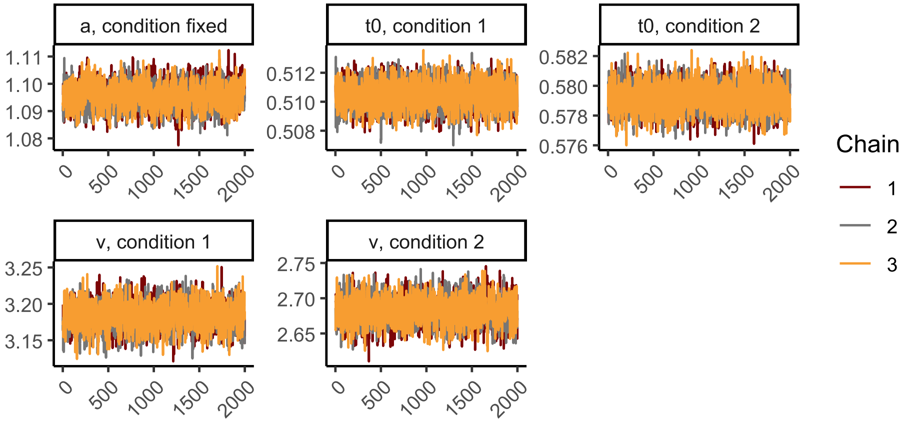



**Figure A1.8.** Convergence of the model in simulation 3. Plots should resemble a 'fat, hairy caterpillar'.

**DDM simulation 3: Two conditions; 10,000 subjects**

We simulated task data for 10,000 participants. All other simulation settings were identical to simulation 2.

The model converged normally (Figure A1.8). Correlations between simulated and recovered parameter estimates were high and very similar to those found in simulation 2, ranging between $r$ = .83 for the drift rate and .95 for the non-decision time (see Figure A1.9). Thus, the benefit of adding more subjects is already saturated around 1,500 participants, with additional participants not improving parameter estimation.

Overall, we conclude that applying hierarchical Bayesian DDM to the ABCD data is feasible.

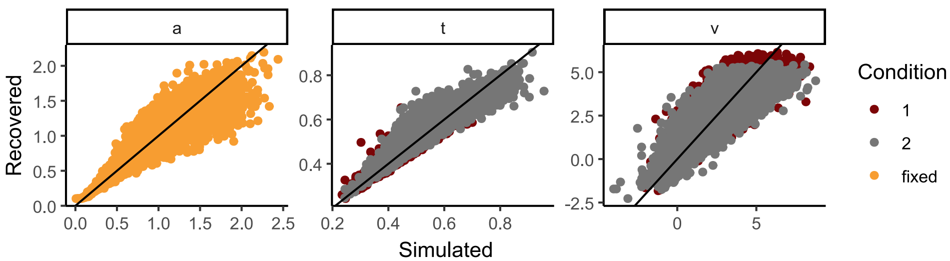



**Figure A1.9.** Parameter recovery in the case of two conditions. a = Boundary Separation; t = Non-Decision Time; v = Drift Rate

## Additional DDM simulations

### DDM simulation 4: Does shrinkage bias the associations between parameter estimates and adversity?

One of the reviewers noted that the hierarchical Bayesian DDM tends to compress parameter estimates by pulling extreme values toward the group mean (a phenomenon known as shrinkage). A concern may be that this could potentially reduce the individual differences of interest, especially if these occur in the tail of the distribution (e.g., the participants with the highest levels of adversity obtaining the most extreme parameter estimates). In general, this is not the case; in contrast, shrinkage tends to pull less reliable and outlier estimates towards the group mean, which has been shown to positively affect the signal-to-noise ratio and reliability of parameter estimates in cognitive neuroscience (Dai et al., 2017; Mejia et al., 2018). To specifically study the effects of shrinkage on the variance of DDM parameter estimates, we nevertheless ran a simulation to investigate the likelihood that shrinkage might obscure adversity-DDM parameter associations.

We simulated DDM parameters for 1,500 participants. Participants' adversity scores followed a log-normal distribution ($mean_{log} = 0$, $sd_{log} = 0.3$) to approximate the skew in the right tail typically found in adversity scores. Drift rates were simulated based on a standardized association of $\beta = 0.1$ with the adversity score. Thus, higher levels of adversity tended to be associated with higher drift rates. Based on the simulated DDM parameters, we simulated 20 trials (RTs and accuracy) per participant, which were then used as input to the DDM model. We used the first 2,000 samples as burn-in, and then took an additional 2,000 samples. We sampled across three chains, which were subsequently combined, for a total of 6,000 samples.

All parameters were recovered with high correlations ranging between 0.84 and 0.97. Figure A1.10 shows signs of shrinkage, especially in the right tail of the drift rate distribution. However, the difference in standard deviations was minimal ($SD_{simulated}$ = 1.52; $SD_{recovered}$ = 1.46.

Next, we calculated the deviations between each simulated and recovered parameter estimate and plotted this against the adversity scores (See Figure A1.11). None of the associations were statistically significant (all *p*s > .05 for linear and quadratic effects).

Finally, we fitted a linear mixed model predicting drift rates estimates as a function of adversity, dataset (simulated vs. recovered; dummy-coded with simulated as the reference category), and the adversity x dataset interaction to assess whether the difference between simulated and recovered drift rates would be different at low, average, and high levels of adversity. We did not find a significant adversity x dataset interaction, $b = -0.03$, $p = .163$. As Figure A1.12 illustrates, there seemed to be small shrinkage ef-

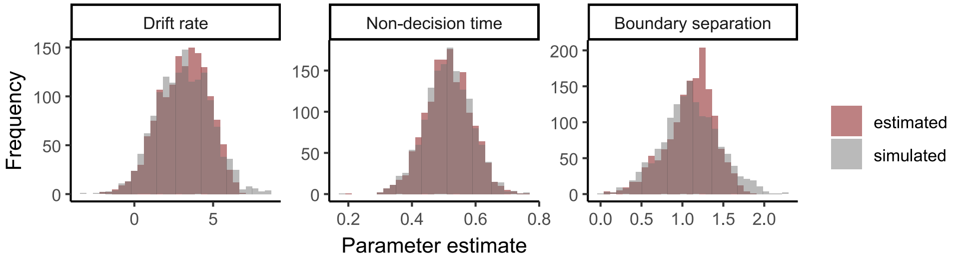



**Figure A1.10.** Histograms of simulated and recovered parameter estimates. a = Boundary Separation; t = Non-Decision Time; v = Drift Rate.

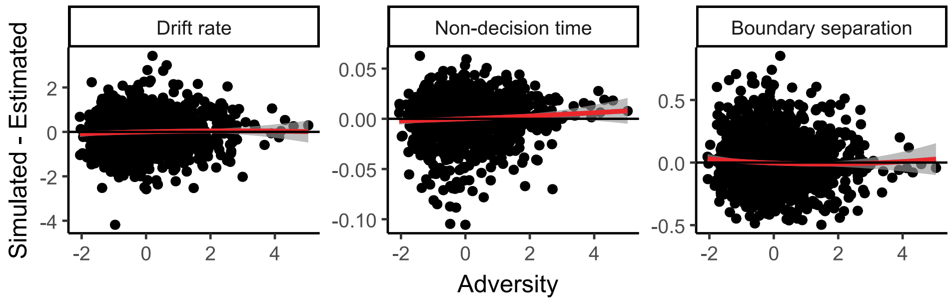



**Figure A1.11.** Deviation between simulated and recovered parameter estimates as a function of adversity. The regression lines show quadratic effects. a = Boundary Separation; t = Non-Decision Time; v = Drift Rate.

fects at the low and high levels of adversity. However, none of these simple slope effects were statistically significant. Taken together, we conclude that DDM recovery at higher levels of adversity was not less precise compared to lower levels of adversity.

### DDM simulation 5: Imputation of missing RTs

To demonstrate the feasibility of the imputation approach for the Mental Rotation Task, we ran a simulation based on 1500 participants in which RT and accuracy data were generated modeled on the real Mental Rotation Task data (RT: $M_{real}$ = 2.65, $M_{sim}$ = 2.76; Accuracy: $M_{real}$ = 59.25%, $M_{sim}$ = 67.23%; RTs above 5 s cut-off: $M_{real}$ = 10.04%, $M_{sim}$ = 8.18%). We fitted two DDM models: one that was fit to the complete data (including RTs > 5 s) and one that was fit to data in which all RTs > 5 s were set to missing. In the latter case, missing RTs were imputed as described above. All other model fit settings were identical to simulations 2-4. Correlations between DDM parameters based on the complete data and imputed data were near perfect, $r$ = 1 for drift rate, $r$ = 0.996 for non-decision time, and $r$ = 0.993 for boundary separation.

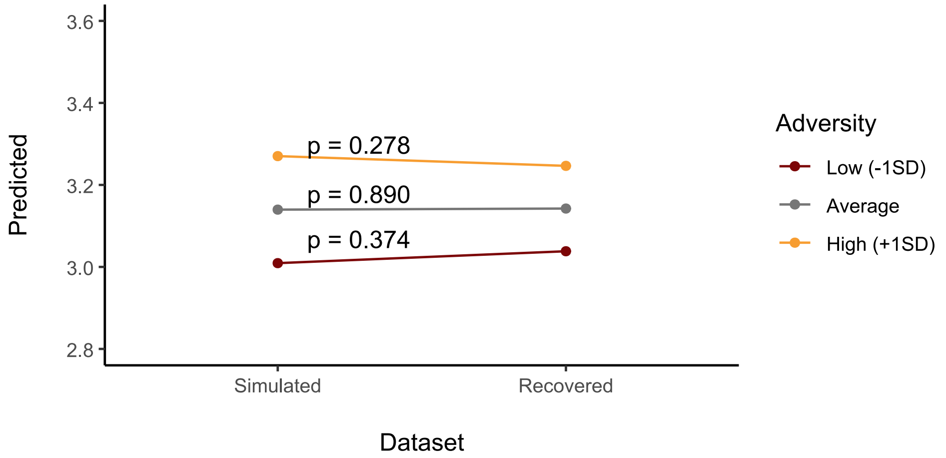



**Figure A1.12.** Simple slopes of the difference between simulated and estimated drift rates at different levels of adversity.

## DDM Model Fit Assessments

### Parameter recovery

The results of our parameter recovery analyses are summarised in Table A1.3 (preregistered approach) and Table A1.2 (updated approach). Using the preregistered approach (simulating the same number of trials as the real data), four out of 16 correlations fell below the pre-specified cut-off (See Table A1.3). Specifically, this was the case for four out of 16 correlations: accuracies for Flanker (.79), Attention Shifting (.73), Processing Speed (.65), and the 75th percentile of RTs for Mental Rotation (.76). In an updated procedure, we increased the number of simulated trials to 100 per task. In these analyses, all correlations were above the .80 cut-off.

Thus, the updated simulation procedure was almost identical to the preregistered procedure, outlined above. The only difference concerned the number of simulated trials. In research with adult participants, it is standard to match the number of simulated trials to the number of observed trials (Lewandowsky & Farrell, 2010). This rule of thumb is arbitrary; researchers sometimes simulate thousands of trials in dedicated parameter recovery studies. The only reason why they typically do not is because it is often sufficient to match the number of observed trials. In our preregistered plan, we followed the convention by matching the number of simulated trials to the number of observed trials. However, in the adult literature, participants frequently complete several dozen, if not hundreds, of trials. In hindsight, we did not sufficiently reflect on this difference, given the lower number of trials per participant in this study involving children. That is, while we followed the convention in the field, the number of simulated trials was lower than is typically the case (because our study involved children), and would also be very low in adult samples.

If the youth's DDM parameters were not recovered accurately because the data were too sparse, increasing the number of simulated trials should not improve these correlations. In other words, if DDM parameters contained a lot of measurement noise or were biased, the correlation between real and simulated RTs/accuracy would remain low even if we simulated more trials. However, that is not what we found. Instead, all correlations were above the .80 cut-off. Many even surpassed .90. This indicated that our data quality was good—the lower correlations observed in the preregistered analysis were solely due to the low number of simulated trials—and that we successfully recovered DDM parameters once addressing this issue.

**Table A1.2.** Simulation-based model fit assessment comparing observed and predicted data using 100 simulated trials (accuracy, 25th, 50th, 75th percentile).

| Task | Condition | 25th Percentile | 50th Percentile | 75th Percentile | Accuracy | R^ |
|---|---|---|---|---|---|---|
| Flanker - Model 1 | congruent | 0.96 | 0.96 | 0.95 | 0.53 | 1.008 |
| Flanker - Model 1 | incongruent | 0.94 | 0.95 | 0.94 | 0.93 | 1.008 |
| **Flanker - Model 2** | | **0.96** | **0.96** | **0.96** | **0.90** | **1.011** |
| **Mental Rotation - Model 1** | | **0.90** | **0.88** | **0.84** | **0.94** | **1.010** |
| Mental Rotation - Model 2 | | 0.87 | 0.88 | 0.86 | 0.96 | 1.010 |
| Attention Shifting - Model 1 | repeat | 0.80 | 0.82 | 0.80 | -0.00 | 1.009 |
| Attention Shifting - Model 1 | switch | 0.82 | 0.82 | 0.83 | 0.46 | 1.009 |
| **Attention Shifting - Model 2** | | **0.94** | **0.95** | **0.95** | **0.88** | **1.008** |
| Processing Speed - Model 1 | | 0.92 | 0.94 | 0.93 | 0.81 | 1.013 |
| **Processing Speed - Model 2** | | **0.94** | **0.94** | **0.93** | **0.80** | **1.011** |

*Note:* The models that were selected for inclusion in the primary analyses are printed in bold.

**Table A1.3.** Simulation-based model fit assessment comparing observed and predicted data using the same number of observed and simulated trials (accuracy, 25th, 50th, 75th percentile).

| Task | Condition | 25th Percentile | 50th Percentile | 75th Percentile | Accuracy | R^ |
|---|---|---|---|---|---|---|
| Flanker - Model 1 | congruent | 0.88 | 0.88 | 0.87 | 0.30 | 1.008 |
| Flanker - Model 1 | incongruent | 0.88 | 0.89 | 0.88 | 0.88 | 1.008 |
| **Flanker - Model 2** | | **0.91** | **0.92** | **0.93** | **0.79** | **1.011** |
| **Mental Rotation - Model 1** | | **0.84** | **0.80** | **0.76** | **0.87** | **1.010** |
| Attention Shifting - Model 1 | repeat | 0.72 | 0.74 | 0.73 | -0.01 | 1.009 |
| Attention Shifting - Model 1 | switch | 0.69 | 0.67 | 0.68 | 0.26 | 1.009 |
| **Attention Shifting - Model 2** | | **0.91** | **0.91** | **0.90** | **0.73** | **1.008** |
| Processing Speed - Model 1 | | 0.88 | 0.89 | 0.88 | 0.66 | 1.013 |
| **Processing Speed - Model 2** | | **0.90** | **0.89** | **0.88** | **0.65** | **1.011** |

*Note:* The models that were selected for inclusion in the primary analyses are printed in bold.

As planned, we explored whether model fit would be relatively worse at different levels of the two measures of adversity. Table A1.4 and A1.5 present correlations between observed and predicted RTs and accuracy at different levels of adversity. Model fit was high for all tasks across all levels of adversity, and there were no indications for any meaningful differences.

**Table A1.4.** Simulation-based model fit assessment at different levels of material deprivation comparing observed and predicted data using 100 simulated trials (accuracy, 25th, 50th, 75th percentile).

| Task | Material deprivation | 25th Percentile | 50th Percentile | 75th Percentile | Accuracy |
|---|---|---|---|---|---|
| Flanker - Model 2 | < -1SD | 0.97 | 0.97 | 0.96 | 0.82 |
| Flanker - Model 2 | > 1SD | 0.95 | 0.95 | 0.96 | 0.93 |
| Flanker - Model 2 | -1SD ≥ x ≤ 1SD | 0.96 | 0.96 | 0.96 | 0.87 |
| Mental Rotation - Model 1 | < -1SD | 0.88 | 0.86 | 0.82 | 0.94 |
| Mental Rotation - Model 1 | -1SD ≥ x ≤ 1SD | 0.89 | 0.88 | 0.84 | 0.94 |
| Mental Rotation - Model 1 | > 1SD | 0.92 | 0.89 | 0.85 | 0.93 |
| Attention Shifting - Model 2 | < -1SD | 0.94 | 0.95 | 0.95 | 0.85 |
| Attention Shifting - Model 2 | > 1SD | 0.93 | 0.94 | 0.94 | 0.90 |
| Attention Shifting - Model 2 | -1SD ≥ x ≤ 1SD | 0.94 | 0.95 | 0.95 | 0.86 |
| Processing Speed - Model 2 | < -1SD | 0.95 | 0.96 | 0.94 | 0.82 |
| Processing Speed - Model 2 | -1SD ≥ x ≤ 1SD | 0.94 | 0.94 | 0.93 | 0.79 |
| Processing Speed - Model 2 | > 1SD | 0.93 | 0.93 | 0.93 | 0.81 |

**Table A1.5.** Simulation-based model fit assessment at different levels of household threat comparing observed and predicted data using 100 simulated trials (accuracy, 25th, 50th, 75th percentile).

| Task | Household threat | 25th Percentile | 50th Percentile | 75th Percentile | Accuracy |
|---|---|---|---|---|---|
| Flanker - Model 2 | < -1SD | 0.96 | 0.96 | 0.96 | 0.83 |
| Flanker - Model 2 | > 1SD | 0.96 | 0.96 | 0.95 | 0.92 |
| Flanker - Model 2 | -1SD ≥ x ≤ 1SD | 0.96 | 0.96 | 0.96 | 0.88 |
| Mental Rotation - Model 1 | < -1SD | 0.88 | 0.87 | 0.82 | 0.94 |
| Mental Rotation - Model 1 | -1SD ≥ x ≤ 1SD | 0.90 | 0.88 | 0.84 | 0.94 |
| Mental Rotation - Model 1 | > 1SD | 0.90 | 0.88 | 0.85 | 0.94 |
| Attention Shifting - Model 2 | < -1SD | 0.95 | 0.95 | 0.95 | 0.84 |
| Attention Shifting - Model 2 | > 1SD | 0.94 | 0.94 | 0.94 | 0.86 |
| Attention Shifting - Model 2 | -1SD ≥ x ≤ 1SD | 0.94 | 0.95 | 0.95 | 0.88 |
| Processing Speed - Model 2 | < -1SD | 0.94 | 0.94 | 0.94 | 0.79 |
| Processing Speed - Model 2 | -1SD ≥ x ≤ 1SD | 0.94 | 0.94 | 0.93 | 0.80 |
| Processing Speed - Model 2 | > 1SD | 0.93 | 0.93 | 0.93 | 0.81 |

For the Processing Speed Task, we found a high degree of Kurtosis in the left-hand tail of the non-decision time distribution. This tail consisted of participants with the lowest RTs (between 0.3s and ~1s). Although overall accuracy on the Processing Speed Task was very high (96.41%), overall accuracy for RTs < 1s was below chance, with accuracy increasing above change starting at 1s. Therefore, we decided to remove RTs < 1s (0.1% of trials) and refit the model, which solved the kurtosis in non-decision times. Recovery of RTs was above the cut-off of .80 for each quantile. Thus, we selected this model for the main analyses.

### Model convergence

Figure A1.13-A1.16 show model convergence for each task. All models converged normally.

### Parameter distributions

Figure A1.17-A1.20 show the distributions of DDM parameters for each task.

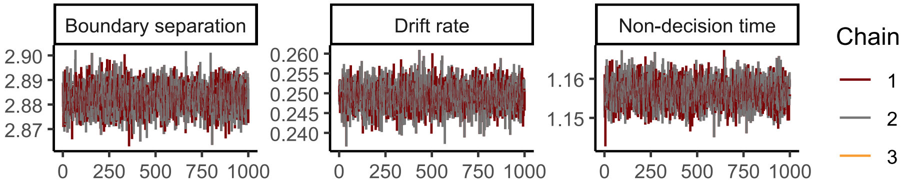



**Figure A1.13.** Convergence of the final model for the Mental Rotation Task. Plots should resemble a 'fat, hairy caterpillar'.

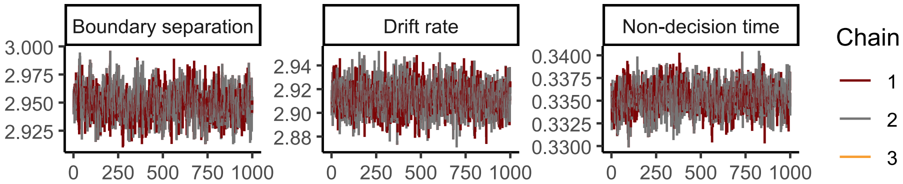



**Figure A1.14.** Convergence of the final model for the Inhibition Task. Plots should resemble a 'fat, hairy caterpillar'.

## SEM Fit

The factor loadings and residual variances of the full test model are presented in Table A1.6. Table A1.7 presents the correlations between latent variables in the model.

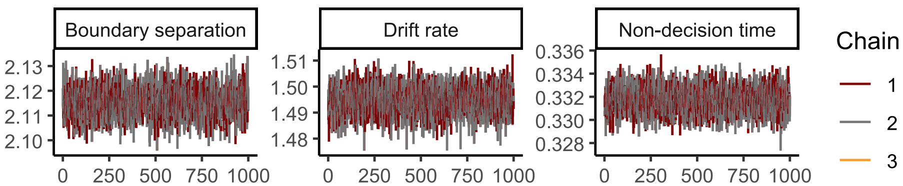



**Figure A1.15.** Convergence of the final model for the Attention Shifting Task. Plots should resemble a 'fat, hairy caterpillar'.

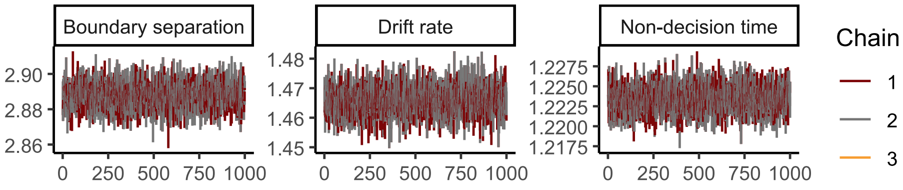



**Figure A1.16.** Convergence of the final model for the Processing Speed Task. Plots should resemble a 'fat, hairy caterpillar'.

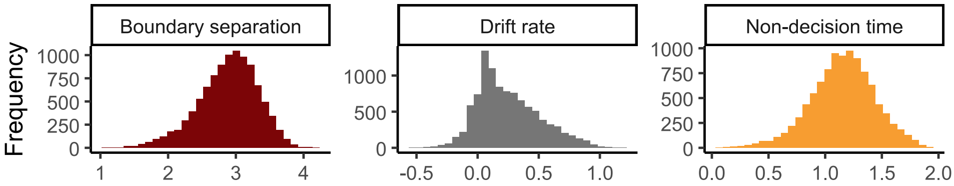



**Figure A1.17.** Parameter distributions in the final model of the Mental Rotation Task.

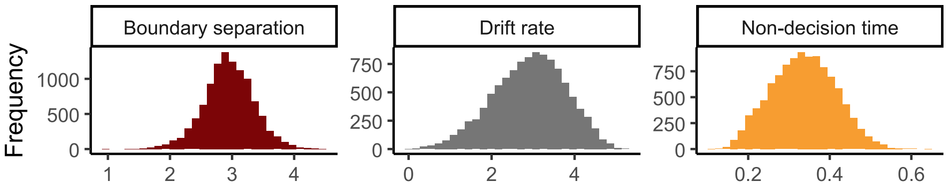



**Figure A1.18.** Parameter distributions in the final model of the Inhibition Task.

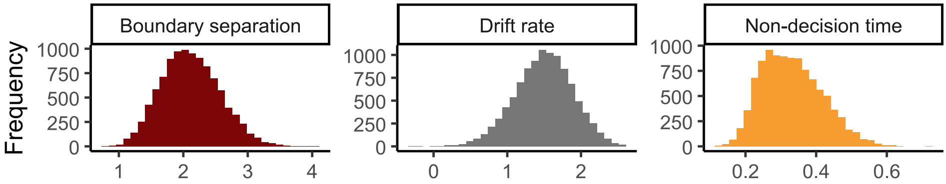



**Figure A1.19.** Parameter distributions in the final model of the Attention Shifting Task.

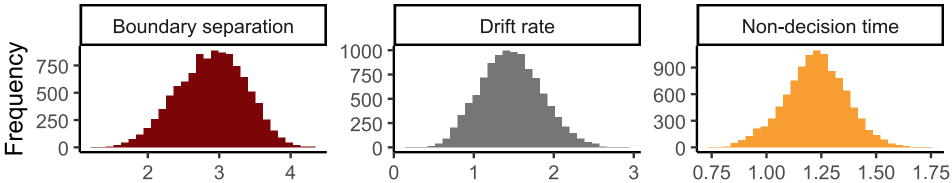



**Figure A1.20.** Parameter distributions in the final model of the Processing Speed Task.

**Table A1.6.** Factor loadings and unstandardized residual variances in the test set.

| | Estimate (unstandardized) | SE | Z | p | Estimate standardized |
|---|---|---|---|---|---|
| **Factor loadings** | | | | | |
| **Task-general drift rate** | | | | | |
| Processing Speed Task | 1.00 | 0.00 | | | 0.52 |
| Attention Shifting Task | 1.19 | 0.04 | 29.43 | 0.000 | 0.63 |
| Mental Rotation Task | 0.51 | 0.03 | 18.37 | 0.000 | 0.27 |
| Inhibition Task | 1.22 | 0.04 | 29.76 | 0.000 | 0.65 |
| **Task-general boundary separation** | | | | | |
| Processing Speed Task | 1.00 | 0.00 | | | 0.55 |
| Attention Shifting Task | 1.44 | 0.04 | 38.19 | 0.000 | 0.80 |
| Mental Rotation Task | 0.28 | 0.02 | 11.97 | 0.000 | 0.15 |
| Inhibition Task | 1.14 | 0.03 | 37.63 | 0.000 | 0.63 |
| **Task-general non-decision time** | | | | | |
| Processing Speed Task | 1.00 | 0.00 | | | 0.45 |
| Attention Shifting Task | 1.46 | 0.05 | 27.28 | 0.000 | 0.66 |
| Mental Rotation Task | 0.67 | 0.03 | 19.49 | 0.000 | 0.30 |
| Inhibition Task | 1.53 | 0.05 | 29.78 | 0.000 | 0.70 |
| **Residual variances** | | | | | |
| **Task-specific drift rate** | | | | | |
| Inhibition Task | 0.55 | 0.01 | 43.01 | 0.000 | |
| Attention Shifting Task | 0.52 | 0.01 | 42.61 | 0.000 | |
| Mental Rotation Task | 0.82 | 0.01 | 64.81 | 0.000 | |
| Processing Speed Task | 0.71 | 0.01 | 55.58 | 0.000 | |
| **Task-specific boundary separation** | | | | | |
| Inhibition Task | 0.61 | 0.01 | 53.53 | 0.000 | |
| Attention Shifting Task | 0.39 | 0.01 | 31.86 | 0.000 | |
| Mental Rotation Task | 0.95 | 0.01 | 66.89 | 0.000 | |
| Processing Speed Task | 0.69 | 0.01 | 58.55 | 0.000 | |
| **Task-specific non-decision time** | | | | | |
| Inhibition Task | 0.56 | 0.01 | 44.13 | 0.000 | |
| Attention Shifting Task | 0.61 | 0.01 | 44.01 | 0.000 | |
| Mental Rotation Task | 0.86 | 0.01 | 64.90 | 0.000 | |
| Processing Speed Task | 0.78 | 0.01 | 60.72 | 0.000 | |

**Table A1.7.** Correlations between latent task-general and task-specific factors in the test set.

| | Correlation |
|---|---|
| **Task-general** | |
| Drift rate - Boundary separation | -0.575*** |
| Drift rate - Non-decision time | -0.046* |
| Boundary separation - Non-decision time | 0.710*** |
| **Task-specific Inhibition Task** | |
| Drift rate - Boundary separation | -0.098*** |
| Drift rate - Non-decision time | 0.019 |
| Boundary separation - Non-decision time | 0.340*** |
| **Task-specific Attention Shifting Task** | |
| Drift rate - Boundary separation | -0.106*** |
| Drift rate - Non-decision time | 0.030* |
| Boundary separation - Non-decision time | -0.228*** |
| **Task-specific Mental Rotation Task** | |
| Drift rate - Boundary separation | 0.305*** |
| Drift rate - Non-decision time | 0.230*** |
| Boundary separation - Non-decision time | 0.102*** |
| **Task-specific Processing Speed Task** | |
| Drift rate - Boundary separation | -0.125*** |
| Drift rate - Non-decision time | 0.052*** |
| Boundary separation - Non-decision time | -0.097*** |

# Appendix 2 - Chapter 3

## Section 1. Descriptive statistics of adversity measures

Figures A2.1-A2.3 present histograms of each separate adversity measure, as well as the composite adversity measures used in the analyses. See the main text for more information on how the composites were calculated.

## Section 2. Additional information on cognitive tasks

### Distributions of response times and error rates

Figure A2.4 and A2.5 show the distributions of mean response times and mean error rate for each cognitive task.

### Condition manipulation checks

Table A2.1 presents manipulation checks for response times for the inhibition and attention-shifting tasks. The manipulation checks were based on mean log-transformed response times using paired-sample *t*-tests. For each task, we tested whether there was a significant difference in mean log-transformed response time between the congruent (repeat) condition and the incongruent (switch) condition. All tasks showed a significant difference in the expected direction, with participants on average being faster on the congruent (repeat) condition compared to the incongruent (switch) condition (all $ps < .001$).

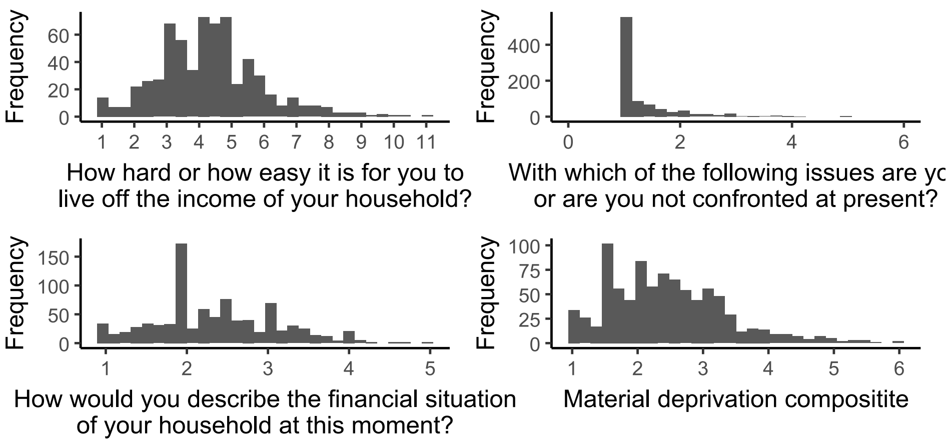



**Figure A2.1.** Histograms of material deprivation measures

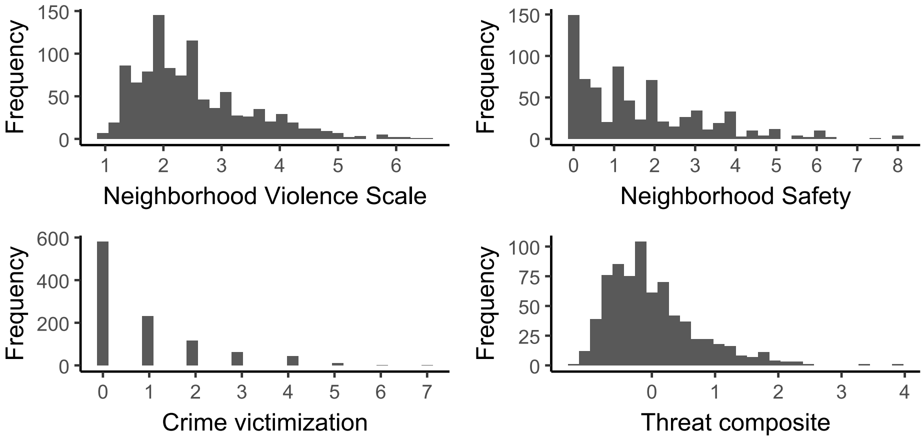



**Figure A2.2.** Histograms of threat measures

**Table A2.1.** Log-transformed response time differences across conditions for each task.

| Task | Estimate | t | p |
|---|---|---|---|
| Flanker task | 0.14 | 51.96 | < .001 |
| Simon task | 0.02 | 7.27 | < .001 |
| Color-shape task | 0.07 | 19.58 | < .001 |
| Global-local task | 0.04 | 8.86 | < .001 |
| Animacy-size task | 0.16 | 40.65 | < .001 |

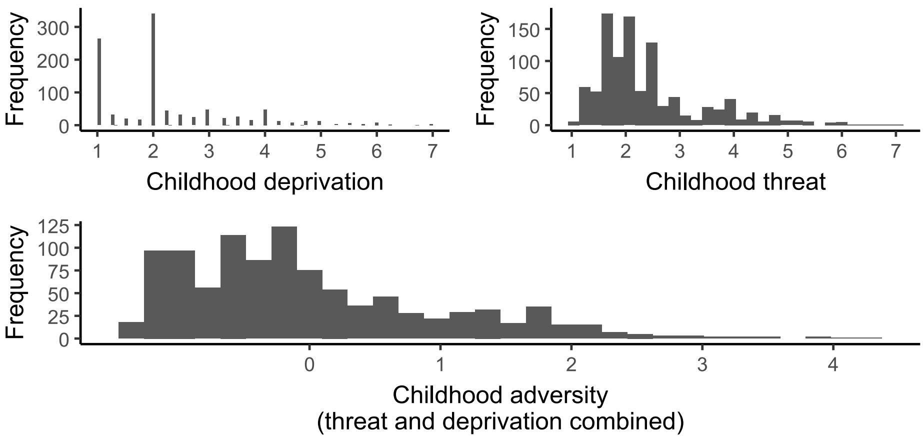



**Figure A2.3.** Histograms of threat measures

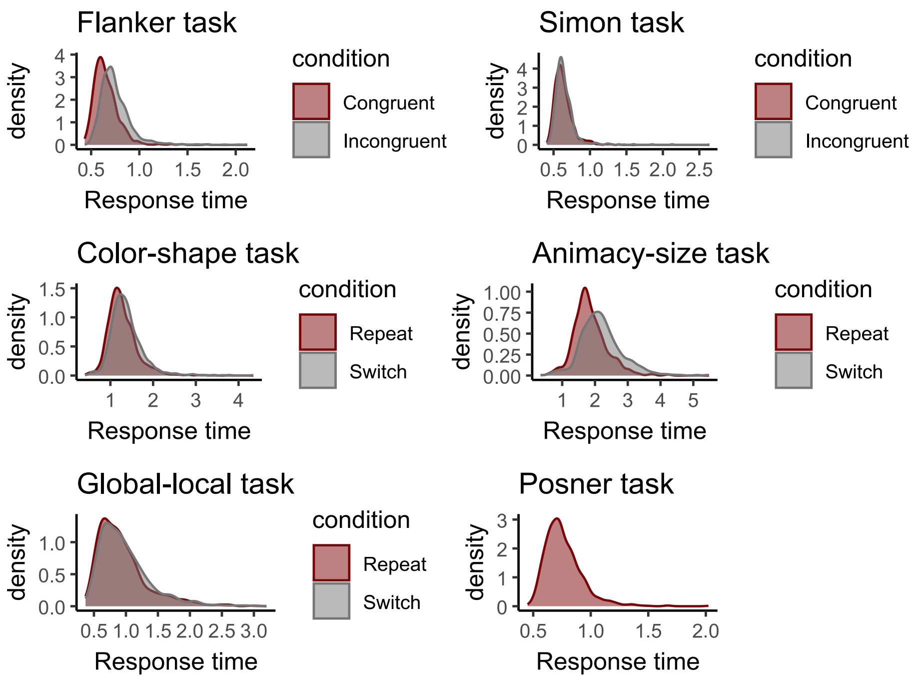



**Figure A2.4.** Distributions of response times for all tasks.

## Split-half reliability

For each cognitive task, we calculated split-half reliabilities using the `splithalf` package (Parsons, 2021). We calculated split-half reliabilities for mean response times and error rates, separately for each condition (see Table A2.2 and Table A2.3). The split-half reliability of mean response times was high across all tasks and conditions. For error rates, the reliability estimates were generally lower, which is likely due to ceiling effects (see also Figure A2.5).

**Table A2.2.** Split-half reliabilities of mean response times for all cognitive tasks.

| | Split-half reliability | | Spearman-Brown corrected | |
|---|---|---|---|---|
| Task | Congruent/Repeat | Incongruent/Switch | Congruent/Repeat | Incongruent/Switch |
| Flanker task | 0.91 [0.86, 0.93] | 0.90 [0.87, 0.93] | 0.95 [0.93, 0.96] | 0.95 [0.93, 0.96] |
| Simon task | 0.88 [0.83, 0.92] | 0.89 [0.83, 0.92] | 0.94 [0.91, 0.96] | 0.94 [0.90, 0.96] |
| Color-shape task | 0.83 [0.81, 0.85] | 0.86 [0.84, 0.87] | 0.91 [0.89, 0.92] | 0.92 [0.91, 0.93] |
| Animacy-size task | 0.82 [0.79, 0.84] | 0.86 [0.85, 0.88] | 0.90 [0.89, 0.91] | 0.93 [0.92, 0.93] |
| Global-local task | 0.82 [0.79, 0.85] | 0.82 [0.78, 0.86] | 0.90 [0.88, 0.92] | 0.90 [0.88, 0.92] |
| Posner task | 0.95 [0.94, 0.95] | | 0.97 [0.97, 0.98] | |

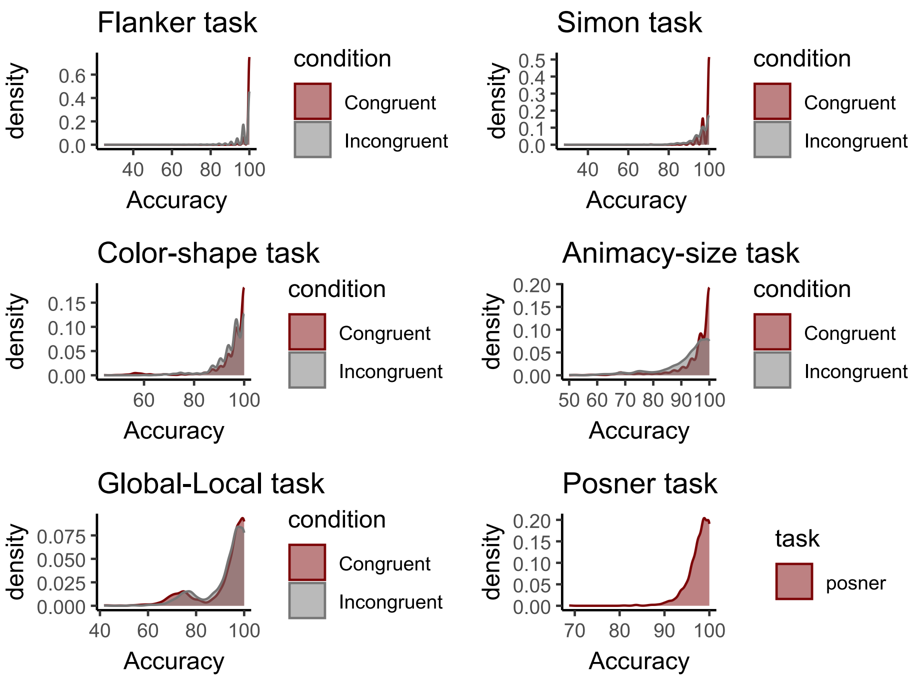



**Figure A2.5.** Distributions of error rates for all tasks.

**Table A2.3.** Split-half reliabilities of error rates for all cognitive tasks.

| | Split-half reliability | | Spearman-Brown corrected | |
|---|---|---|---|---|
| Task | Congruent/Repeat | Incongruent/Switch | Congruent/Repeat | Incongruent/Switch |
| Flanker task | 0.57 [0.40, 0.65] | 0.64 [0.60, 0.69] | 0.72 [0.58, 0.79] | 0.78 [0.75, 0.81] |
| Simon task | 0.52 [0.46, 0.57] | 0.48 [0.44, 0.53] | 0.68 [0.63, 0.72] | 0.65 [0.61, 0.69] |
| Color-shape task | 0.74 [0.70, 0.77] | 0.46 [0.41, 0.50] | 0.85 [0.82, 0.87] | 0.63 [0.58, 0.67] |
| Animacy-size task | 0.67 [0.62, 0.71] | 0.59 [0.55, 0.63] | 0.80 [0.77, 0.83] | 0.74 [0.71, 0.77] |
| Global-local task | 0.67 [0.63, 0.71] | 0.58 [0.54, 0.62] | 0.80 [0.77, 0.83] | 0.73 [0.70, 0.77] |
| Posner task | 0.42 [0.37, 0.46] | | 0.59 [0.54, 0.63] | |

## Section 3. Drift Diffusion Modeling

### Model convergence

Figures A2.6-A2.11 present the convergence of MCMC chains of the Hierarchical Drift Diffusion Models for all tasks. The figures should resemble a “fat, hairy caterpillar”, which was the case for all tasks.

In addition, we calculated the R^ statistic (also known as the Gelman-Rubin statistic) (Gelman & Rubin, 1992). The R^ calculates the ratio between the variation be-

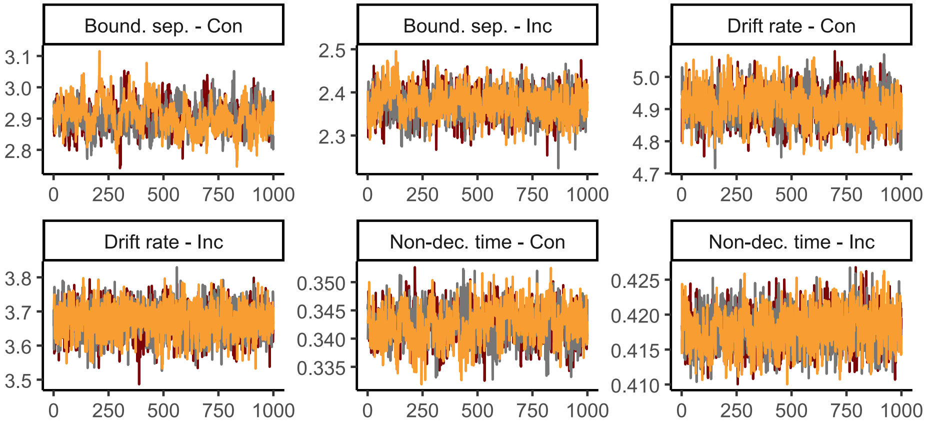



**Figure A2.6.** Trace convergence across three chains for the Flanker Task.

tween MCMC chains and the variation within MCMC chains. A general guideline is that R^ values should be smaller than 1.1. All R^ values are presented in Table A2.4 below.

**Simulation-based model fit assessment**

Table A2.4 presents simulation-based model fit assessments for all tasks. The simulation procedure was as follows. First, we used the DDM parameter estimates for each participant and used them to simulate response times and accuracy data (100 trials per participant). Then, we computed correlations between the simulated and observed

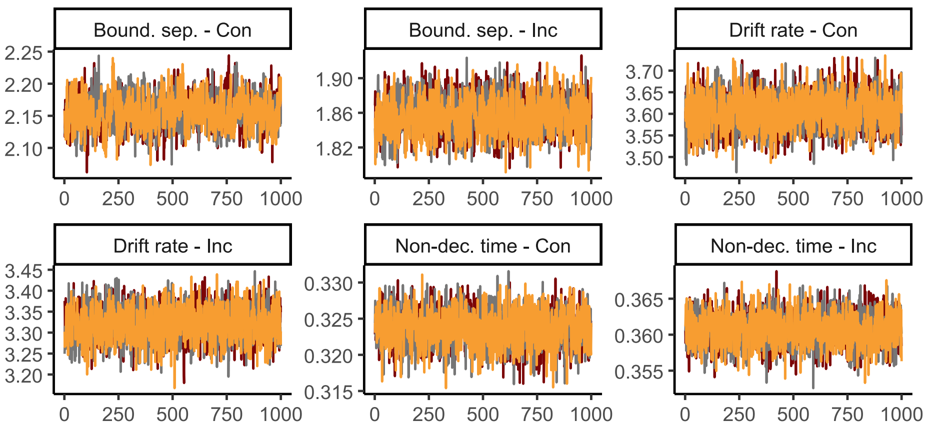



**Figure A2.7.** Trace convergence across three chains for the Simon Task.

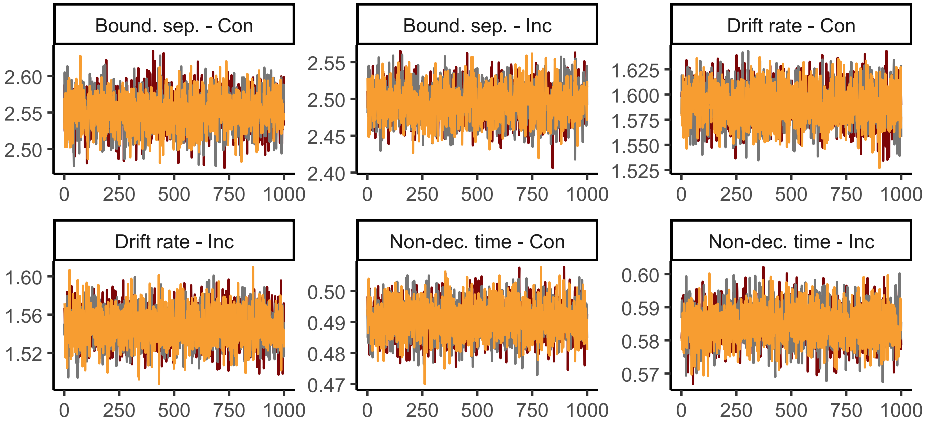



**Figure A2.8.** Trace convergence across three chains for the Color-shape Task.

response times and accuracies. In the case of response times, we did so separately at the 25th, 50th, and 75th percentile. In the case of accuracy, we looked at mean accuracy. All correlations were > .89, indicating good model fit.

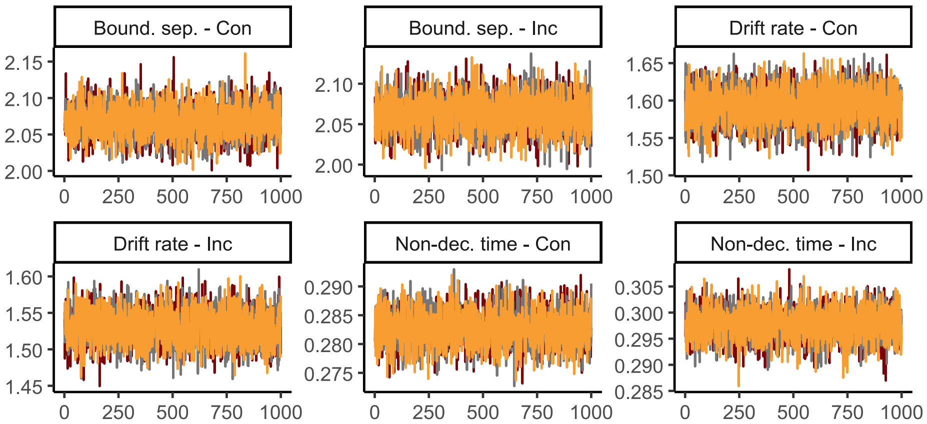



**Figure A2.9.** Trace convergence across three chains for the Global-local Task.

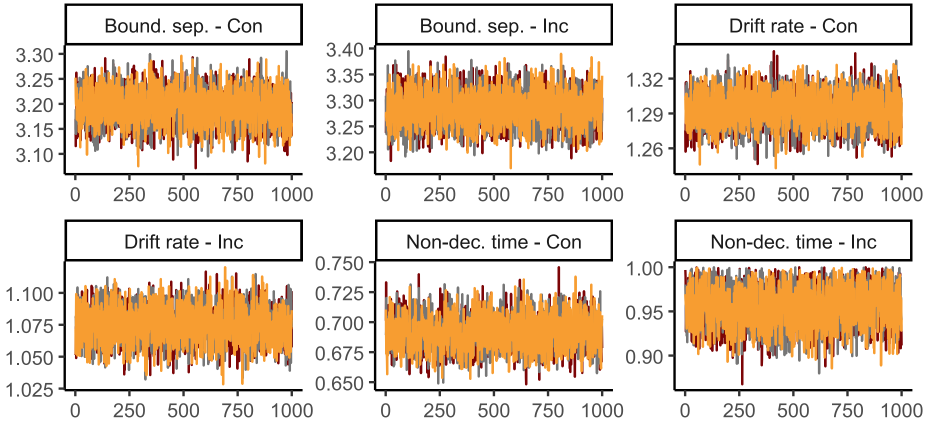



**Figure A2.10.** Trace convergence across three chains for the Animacy-size Task.

**Table A2.4.** Model fit assessment

| Task | Condition | RT - 25th Percentile | RT - 50th Percentile | RT - 75th Percentile | Accuracy | R^ |
|---|---|---|---|---|---|---|
| Flanker task | Congruent | 0.98 | 0.98 | 0.97 | 0.89 | 1.011 |
| Flanker task | Incongruent | 0.97 | 0.97 | 0.98 | 0.94 | 1.011 |
| Simon task | Congruent | 0.97 | 0.97 | 0.97 | 0.91 | 1.003 |
| Simon task | Incongruent | 0.95 | 0.97 | 0.97 | 0.92 | 1.003 |
| Color-shape task | Repeat | 0.97 | 0.96 | 0.97 | 0.98 | 1.006 |
| Color-shape task | Switch | 0.96 | 0.97 | 0.97 | 0.95 | 1.006 |
| Animacy-size task | Repeat | 0.96 | 0.96 | 0.96 | 0.98 | 1.002 |
| Animacy-size task | Switch | 0.95 | 0.96 | 0.96 | 0.97 | 1.002 |
| Global-local task | Repeat | 0.95 | 0.96 | 0.96 | 0.98 | 1.001 |
| Global-local task | Switch | 0.95 | 0.96 | 0.96 | 0.97 | 1.001 |
| Posner task | | 0.97 | 0.98 | 0.98 | 0.89 | 1.002 |

*Note:* Simulation-based model fit assessment compared observed and predicted data using 100 simulated trials (accuracy, 25th, 50th, 75th percentile). In addition, We calculated R^ values, which should be below 1.1 to indicate adequate chain convergence.

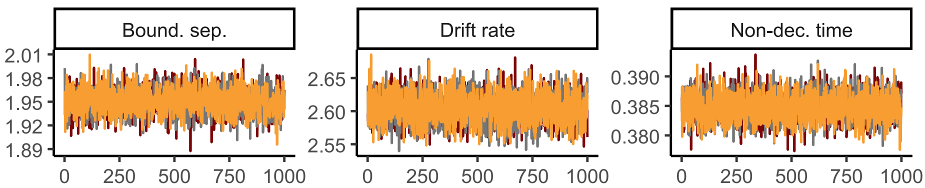



**Figure A2.11.** Trace convergence across three chains for the Posner Task.

**Section 4. Effects of environmental variables**

Table A2.5 presents effects of environmental noise and mean-centered state anxiety on Drift Diffusion parameters. Effects of environmental noise were mostly small non-significant. Mean-centered differences in state anxiety were negatively associated with drift rates across all tasks. In sessions where participants were more anxious than average, their drift rates across all tasks were lower.

**Table A2.5.** Effects of noise and anxiety on Drift Diffusion parameters.

| Task | Intercept | | Noise | | Anxiety | |
|---|---|---|---|---|---|---|
| | Estimate | *p* | Estimate | *p* | Estimate | *p* |
| Animacy-Size (Rep.) - Bound. sep. | 3.23 (0.04) | < .001 | -0.03 (0.02) | .093 | 0.01 (0.02) | .610 |
| Animacy-Size (Sw.) - Bound. sep. | 3.33 (0.05) | < .001 | -0.04 (0.02) | .080 | -0.01 (0.02) | .558 |
| Animacy-Size (Rep.) - Drift rate | 1.34 (0.02) | < .001 | -0.02 (0.01) | .062 | -0.04 (0.01) | < .001 |
| Animacy-Size (Sw.) - Drift rate | 1.11 (0.02) | < .001 | -0.02 (0.01) | .155 | -0.04 (0.01) | < .001 |
| Animacy-Size (Rep.) - Non-dec. time | 0.67 (0.01) | < .001 | -0.02 (0.01) | .001 | -0.01 (0.00) | .267 |
| Animacy-Size (Sw.) - Non-dec. time | 0.80 (0.01) | < .001 | -0.01 (0.01) | .011 | -0.00 (0.01) | .701 |
| Color-Shape (Rep.) - Bound. sep. | 2.57 (0.03) | < .001 | -0.01 (0.02) | .578 | -0.01 (0.02) | .675 |
| Color-Shape (Sw.) - Bound. sep. | 2.53 (0.03) | < .001 | -0.02 (0.02) | .256 | -0.00 (0.02) | .839 |
| Color-Shape (Rep.) - Drift rate | 1.65 (0.03) | < .001 | -0.03 (0.01) | .034 | -0.06 (0.01) | < .001 |
| Color-Shape (Sw.) - Drift rate | 1.59 (0.03) | < .001 | -0.02 (0.01) | .153 | -0.06 (0.01) | < .001 |
| Color-Shape (Rep.) - Non-dec. time | 0.50 (0.01) | < .001 | -0.01 (0.00) | .129 | -0.00 (0.00) | .419 |
| Color-Shape (Sw.) - Non-dec. time | 0.59 (0.01) | < .001 | -0.00 (0.00) | .387 | -0.01 (0.00) | .143 |
| Flanker (Con.) - Bound. sep. | 2.90 (0.03) | < .001 | -0.00 (0.02) | .797 | -0.01 (0.02) | .380 |
| Flanker (Inc.) - Bound. sep. | 2.42 (0.04) | < .001 | -0.03 (0.02) | .121 | -0.04 (0.02) | .045 |
| Flanker (Con.) - Drift rate | 4.92 (0.07) | < .001 | -0.03 (0.03) | .360 | -0.13 (0.03) | < .001 |
| Flanker (Inc.) - Drift rate | 3.77 (0.07) | < .001 | -0.08 (0.03) | .029 | -0.14 (0.03) | < .001 |
| Flanker (Con.) - Non-dec. time | 0.34 (0.00) | < .001 | -0.00 (0.00) | .855 | -0.00 (0.00) | .092 |
| Flanker (Inc.) - Non-dec. time | 0.42 (0.00) | < .001 | -0.00 (0.00) | .898 | 0.00 (0.00) | .767 |
| Global-Local (Rep.) - Bound. sep. | 2.06 (0.04) | < .001 | 0.00 (0.02) | .919 | 0.03 (0.02) | .093 |
| Global-Local (Sw.) - Bound. sep. | 2.04 (0.04) | < .001 | 0.01 (0.02) | .702 | 0.03 (0.02) | .030 |
| Global-Local (Rep.) - Drift rate | 1.71 (0.04) | < .001 | -0.05 (0.02) | .009 | -0.10 (0.02) | < .001 |
| Global-Local (Sw.) - Drift rate | 1.64 (0.04) | < .001 | -0.05 (0.02) | .009 | -0.10 (0.02) | < .001 |
| Global-Local (Rep.) - Non-dec. time | 0.29 (0.00) | < .001 | -0.00 (0.00) | .255 | 0.00 (0.00) | .583 |
| Global-Local (Sw.) - Non-dec. time | 0.30 (0.00) | < .001 | -0.00 (0.00) | .149 | 0.00 (0.00) | .418 |
| Posner - Bound. sep. | 1.98 (0.02) | < .001 | -0.01 (0.01) | .215 | -0.02 (0.01) | .126 |
| Posner - Drift rate | 2.68 (0.04) | < .001 | -0.04 (0.02) | .022 | -0.04 (0.02) | .042 |
| Posner - Non-dec. time | 0.38 (0.00) | < .001 | 0.00 (0.00) | .703 | 0.00 (0.00) | .550 |
| Simon (Con.) - Bound. sep. | 2.15 (0.03) | < .001 | 0.00 (0.01) | .806 | -0.04 (0.01) | .009 |
| Simon (Inc.) - Bound. sep. | 1.88 (0.03) | < .001 | -0.01 (0.01) | .232 | -0.04 (0.01) | .005 |
| Simon (Con.) - Drift rate | 3.66 (0.06) | < .001 | -0.04 (0.03) | .204 | -0.16 (0.03) | < .001 |
| Simon (Inc.) - Drift rate | 3.42 (0.07) | < .001 | -0.06 (0.03) | .061 | -0.19 (0.04) | < .001 |
| Simon (Con.) - Non-dec. time | 0.33 (0.00) | < .001 | -0.00 (0.00) | .087 | -0.00 (0.00) | .101 |
| Simon (Inc.) - Non-dec. time | 0.36 (0.00) | < .001 | -0.00 (0.00) | .986 | -0.00 (0.00) | .044 |

*Note:* State anxiety was mean-centered relative to the overall mean across test sessions. Bound. sep. = boundary separation, Con. = congruent condition, Inc. = incongruent condition, Non-dec. time = non-decision time, Rep. = repeat condition, Sw. = switch condition,

Table A2.6 presents Spearman correlations between preregistered and non-preregistered adversity measures with environmental noise and mean state anxiety. Note that A2.6 includes mean state anxiety, whereas the analyses (and Table A2.5) include the difference from the grand mean in specific testing sessions.

**Table A2.6.** Bivarate correlations between adversity measures, environmental noise, and mean state anxiety.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1. Deprivation in adulthood | - | | | | | |
| 2. Threat in aduldhood | 0.27 | - | | | | |
| 3. Childhood threat | 0.19 | 0.23 | - | | | |
| 4. Childhood deprivation | 0.08 | 0.32 | 0.48 | - | | |
| 5. Environmental noise | 0.05 | 0.17 | 0.07 | 0.14 | - | |
| 6. State anxiety | 0.09 | 0.21 | 0.14 | 0.13 | 0.08 | - |

**Section 5. Indirect effects of confounders**

Table A2.7 summarizes the indirect effects of confounders in the confirmatory models: age, education, sex, childhood adversity, and, in the case of threat as dependent variable, material deprivation. As explained in the section on confounders in the main article, we assume that these confounders are common causes of both adversity exposure in adulthood (the independent variable) and cognitive processes (the dependent variable). This means that the regression coefficients in Table A2.7 and A2.8 should *not* be interpreted as direct effects; rather, they should be interpreted as indirect effects (i.e., the effect at mean levels of adversity exposure).

For both threat and deprivation in adulthood, we found a negative indirect effect of age on task-general drift rate, a positive indirect effect of age on task-general boundary separation, and a positive indirect of age on task-general non-decision time. Thus, older adults with average levels of adversity exposure in adulthood processed information more slowly across tasks, were generally more cautious, and were generally slower at encoding stimuli and/or executing responses. In addition, childhood adversity had a negative indirect effect on task-general drift rate. People with more exposure to childhood adversity, at average levels of adversity exposure in adulthood, processed information more slowly across tasks. Finally, we found a negative indirect effect of education on task-general drift rate, but not on task-general boundary separation or non-decision time. People with a higher completed education, at average levels of adversity exposure in adulthood, processed information faster across tasks. None of the other indirect effects were significant.

**Table A2.7.** Standardized indirect effects of confounders in the confirmatory models.

| | Threat in adulthood | | | | Deprivation in adulthood | | | |
|---|---|---|---|---|---|---|---|---|
| Confounder | β | SE | 95% CI | p | β | SE | 95% CI | p |
| **Task-general drift rate** | | | | | | | | |
| Material deprivation | -0.06 | 0.05 | [-0.16, 0.04] | .241 | | | | |
| Age | -0.12 | 0.04 | [-0.21, -0.04] | .003 | -0.10 | 0.03 | [-0.16, -0.03] | .003 |
| Sex | 0.04 | 0.03 | [-0.01, 0.09] | .130 | 0.01 | 0.02 | [-0.03, 0.05] | .623 |
| Childhood adversity | -0.19 | 0.04 | [-0.27, -0.10] | < .001 | -0.23 | 0.04 | [-0.31, -0.16] | < .001 |
| Education | | | | | 0.14 | 0.04 | [0.07, 0.21] | < .001 |
| **Task-general boundary separation** | | | | | | | | |
| Material deprivation | 0.07 | 0.04 | [-0.02, 0.15] | .137 | | | | |
| Age | 0.37 | 0.04 | [0.29, 0.45] | < .001 | 0.42 | 0.03 | [0.36, 0.48] | < .001 |
| Sex | -0.00 | 0.03 | [-0.05, 0.05] | .854 | -0.00 | 0.02 | [-0.04, 0.04] | .868 |
| Childhood adversity | -0.00 | 0.05 | [-0.09, 0.09] | .994 | -0.03 | 0.04 | [-0.11, 0.04] | .394 |
| Education | | | | | -0.03 | 0.04 | [-0.10, 0.05] | .508 |
| **Task-general non-decision time** | | | | | | | | |
| Material deprivation | 0.01 | 0.04 | [-0.08, 0.10] | .841 | | | | |
| Age | 0.37 | 0.04 | [0.29, 0.45] | < .001 | 0.39 | 0.03 | [0.32, 0.46] | < .001 |
| Sex | 0.01 | 0.03 | [-0.04, 0.06] | .605 | 0.01 | 0.02 | [-0.03, 0.05] | .507 |
| Childhood adversity | -0.04 | 0.05 | [-0.14, 0.05] | .351 | -0.05 | 0.04 | [-0.12, 0.03] | .216 |
| Education | | | | | -0.00 | 0.04 | [-0.08, 0.08] | .923 |

Table A2.8 summarizes the indirect effects of confounders in the exploratory models: sex, and, in the case of childhood threat as dependent variable, childhood material deprivation. Childhood exposure to deprivation had a negative indirect effect on task-general drift rate and task-general boundary separation. People with more exposure to childhood deprivation, at average levels of childhood threat, processed information more slowly across tasks, and were generally more cautious. Sex did not have an indirect effect on either childhood threat or childhood deprivation.

**Table A2.8.** Standardized indirect effects of confounders in the exploratory models of childhood adversity.

| | Childhood threat | | | | Childhood deprivation | | | |
|---|---|---|---|---|---|---|---|---|
| Confounder | β | SE | 95% CI | p | β | SE | 95% CI | p |
| **Task-general drift rate** | | | | | | | | |
| Childhood deprivation | -0.18 | 0.04 | [-0.26, -0.10] | < .001 | | | | |
| Sex | 0.01 | 0.02 | [-0.03, 0.05] | .661 | 0.01 | 0.02 | [-0.03, 0.06] | .579 |
| **Task-general boundary separation** | | | | | | | | |
| Childhood deprivation | 0.12 | 0.04 | [0.04, 0.20] | .002 | | | | |
| Sex | -0.01 | 0.02 | [-0.05, 0.04] | .725 | -0.01 | 0.02 | [-0.05, 0.04] | .770 |
| **Task-general non-decision time** | | | | | | | | |
| Childhood deprivation | 0.06 | 0.04 | [-0.02, 0.14] | .134 | | | | |
| Sex | 0.01 | 0.02 | [-0.03, 0.05] | .629 | 0.01 | 0.02 | [-0.03, 0.05] | .610 |

# Appendix 3 - Chapter 4

## Section 1. Descriptions of exploratory measures

Current state, poverty exposure, impulsivity, future orientation, and depressive symptoms were collected in all three studies. Attentional style was collected in Study 2 only.

### Current state

We assessed state anxiety during the experiment using the state subscale of the State-Trait Anxiety Inventory (STAI-S; Spielberger et al., 1999). The STAI-S contains 20 short items measuring current anxiety (e.g., "I feel tense"). Participants rated each item on a scale of 1 (not at all) to 4 (very much so). An overall state anxiety variable was computed by averaging across the 20 unweighted items.

In addition, participants answered five questions relating to specific states: "Are you currently sick?" (rated as yes or no); "Have you eaten a full meal today?" (rated as yes or no); "How hungry do you feel right now?" (rated from 1 (not at all) to 5 (very hungry)); "How well did you sleep last night?" (rated from 1 (very poorly) to 5 (very well)); "How rested or refreshed did you feel when you woke up this morning?" (rated from 1 (not at all) to 5 (very rested)). We computed an overall sleep deprivation composite by standardizing and averaging across the two unweighted sleep-related items.

### Poverty exposure

Participants' perceived level of resource scarcity before age 13 was measured using seven items (e.g., "Your family had enough money to afford the kind of home you all needed"). Participants rated each item on a scale from 1 (never true) to 5 (very often true). Scores for the first six items were reverse coded so that higher scores indicated more perceived resource scarcity. The items were averaged together to create an unweighted composite scale.

In addition, we measured several indicators of objective SES before age 13. First, participants separately indicated the highest education of their mother and father on an 8-point scale: 'some high school', 'GED', 'high school diploma', 'some college but no college degree', associate's degree', 'bachelor's degree', 'master's degree', or 'doctoral or lab degree'. The mother and father education level were averaged to create an overall unweighted parental education composite. Participants also indicated their family's household income before age 13 on a 6-point scale: 'less than $ 25k/year', '$25k - $49k/year,'$50 - $74k/year', '$75 - $99k/year', '$100 - $149k/year', 'more than $150k/year'. Scores were reverse coded so that higher scores indicated higher levels of poverty.

We created a composite score of poverty exposure before age 13 by averaging together the standardized scores of perceived level of resource scarcity, overall parental education, and household income.

**Impulsivity**

We assessed impulsivity with the Motor Impulsivity subscale of the Barrett Impulsivity Scale (BIS short form; Patton et al., 1995; Spinella, 2007). The Motor Impulsivity subscale of the BIS consists of five items (e.g., "I do things without thinking"). We did not include the Non-planning subscale because it overlapped substantially with the Future Orientation Scale described below. In addition, we did not include the Attention impulsivity subscale because it included items which we deemed to be mostly irrelevant for our target population (e.g., "I 'squirm' at plays or lectures"). We changed the original 4-point rating scale (rarely/never to almost always) to a 5-point rating scale ranging from 1 (never true) to 5 (very often true). An overall impulsivity variable was computed by averaging the five unweighted items.

**Future Orientation**

We assessed future orientation with an adapted version of the Future Orientation Scale (FOS; Steinberg et al., 2009). The original scale consists of 15 sets of opposing items separated by "BUT" (e.g., "Some people like to plan things out one step at a time BUT other people like to jump right into things without planning them out beforehand"). Participants first choose the item that best matches their general preference, and then indicate whether the statement is "really true" or "somewhat true". We adapted this format in a couple of ways. First, we converted the two statements per item to a single statement by picking the statements in the original right-hand column. Second, we adapted the 15 statements from a third-person to a first-person format. These changes were made in an attempt to reduce the cognitive load of the items. We worried that people with less formal education or who were sitting in a noisier environment would struggle with the length of the original items.

In addition, item 8 of the original scale ("[…] other people would rather spend their money right away on something fun than save it for a rainy day") was changed to "I'd rather spend money right away than save it for a rainy day" (i.e., dropping the phrase "on something fun") to make it more general with regard to the thing that money is spent on. For people from adversity, spending money right now instead of saving it for the future might often be born out of necessity (e.g., having just enough money for food and shelter; being in debt) instead of a failure to delay gratification. Finally, the rating scale was adapted from the original 4-point scale (ranging from really true for the left-hand statement to really true for the right-hand statement) to a 5-point scale ranging from 1 (never true) to 5 (very often true). An overall future orientation variable was computed by averaging the 15 unweighted items.

**Depressive symptoms**

We assessed depressive symptoms during the past week using the Center for Epidemiologic Studies Depression Scale (CESD; Radloff, 1977). The scale consists of 20 items (e.g., "I do things without thinking"). Participants rate each item on a scale of 1 (rarely or none of the time (less than 1 day)) to 4 (most or all of the time (5-7 days)). An overall depression variable was computed by averaging the 20 unweighted items.

### Attentional style

We measured attentional style using the Attentional Style Questionnaire (ASQ; Van Calster et al., 2018). The ASQ measures self-reported attentional style, with seven items asking about the participant's propensity for internally oriented attention (e.g., "During an activity, unrelated mmental images and thoughts come to my mind") and seven items about externally oriented attention (e.g., "I am easily drawn to new stimuli (for example, voices of people passing by, as sound in the house, …) that are not relevant to a task I am doing"). Where necessary, items were recoded in such a way that they reflected *distractibility* by internal and external stimuli, respectively, with higher scores reflecting a higher degree of distractibility. We computed unweighted averages separately for internally oriented attention and externally oriented attention.

## Section 2. Exploratory analyses

### Consistency in unpredictability measures

**Pilot.** The EFA yielded five factors based on parallel analysis (see Table A3.1). Based on their contents, we labelled these factors (1) Daily unpredictability; (2) Household routine; (3) Spatial unpredictability; (4) Chaos/clutter; (5) Social unpredictability.

**Table A3.1.** Exploratory factor analysis on unpredictability items in the Pilot (Part 1).

| Item | Household stability/conflict | Monitoring/ neglect | Macro unp. | Disorganization | People in household |
|---|---|---|---|---|---|
| Changes - Family environment. | **0.33** | | | | |
| CHAOS - It was a real zoo in our home. | **0.38** | | | 0.43 | |
| CHAOS - We almost always seemed to be rushed. | **0.43** | | | 0.34 | |
| CHAOS - You couldn't hear yourself think in our home. | **0.43** | | | 0.32 | |
| CHAOS - At home we could talk to each other without being interrupted. | **0.43** | | | | |
| CHAOS - No matter what our family planned, it usually didn't seem to work out. | **0.47** | | | 0.33 | |
| CHAOS - There was very little commotion in our home. | **0.49** | | | | |
| QUIC - My parents had a stable relationship with each other. | **0.51** | | | | 0.39 |
| CHAOS - The atmosphere in our home was calm. | **0.57** | | | | |
| CHAOS - Our home was a good place to relax. | **0.60** | | | | |
| Perceived - My family life was generally inconsistent and unpredictable from day-to-day. | **0.63** | | | | |
| QUIC - I did not feel safe in my home. | **0.63** | | | | |
| CHAOS - There was often a fuss going on at our home. | **0.69** | | | | |
| QUIC - At least one of my parents had punishments that were unpredictable. | **0.69** | | | | |
| Perceived - When I woke up, I often didn't know what could happen in my house that day. | **0.72** | | | | |
| CHAOS - I often got drawn into other people's arguments at home. | **0.72** | | | | |
| QUIC - One of my parents could go from calm to stressed or nervous in an instant. | **0.73** | | | | |
| Perceived - Things were often chaotic in my house. | **0.73** | | | | |
| QUIC - At least one of my parents was unpredictable. | **0.75** | | | | |
| Perceived - I had a hard time knowing what my parent(s) or other people in my house were going to say. | **0.77** | | | | |
| Perceived - My parent(s) frequently had arguments or fights with each other or other people in my childhood. | **0.78** | | | | |
| QUIC - For at least one of my parents, when they were upset I did not know how they would act. | **0.82** | | | | |
| QUIC - One of my parents could go from calm to furious in an instant. | **0.82** | | | | |
| Perceived - My family environment was often tense and "on edge". | **0.94** | | | | |
| QUIC - I usually knew when my parents were going to be home. | | **0.34** | | | |
| QUIC - I had a set morning routine on school days (i.e., I usually did the same thing each day to get ready). | | **0.44** | | | |
| CHAOS - First thing in the day, we had a regular routine at home. | | **0.52** | | | |
| QUIC - My family ate a meal together most days. | | **0.53** | | | |
| QUIC - My family had holiday traditions that we did every year (e.g., cooking a special food at a particular time of year/decorate the house the same way). | | **0.56** | | | |
| QUIC - My family planned activities to do together. | | **0.57** | | | |
| QUIC - My parents kept track of what I ate (e.g., made sure that I didn't skip meals or tried to make sure I ate healthy food). | | **0.65** | | | |
| QUIC - In my afterschool or free time hours at least one of my parents knew what I was doing. | | **0.66** | | | |
| QUIC - I had a bedtime routine (e.g, my parents tucked me in, my parents read me a book, I took a bath). | | **0.70** | | | |
| QUIC - At least one parent made time each day to see how I was doing. | | **0.71** | | | |
| QUIC - At least one of my parents regularly checked that I did my homework. | | **0.72** | | | |
| QUIC - My parents tried to make sure I got a good night's sleep (e.g., I had a regular bedtime, my parents checked to make sure I went to sleep). | | **0.76** | | | |
| QUIC - At least one of my parents regularly kept track of my school progress. | | **0.80** | | | |

**Table A3.2.** Exploratory factor analysis on unpredictability items in the Pilot (Part 2).

| Item | Household stability/conflict | Monitoring/ neglect | Macro unp. | Disorganization | People in household |
|---|---|---|---|---|---|
| Changes - Economic status. | | | **0.39** | | |
| Residential changes | | | **0.69** | | |
| QUIC - I changed schools mid-year. | | | **0.70** | | |
| QUIC - I moved homes. | | | **0.76** | | |
| Changes - Your childhood neighborhood environment. | | | **0.77** | | |
| Changes - Your childhood school environment. | | | **0.81** | | |
| QUIC - I changed schools. | | | **0.87** | | |
| QUIC - I worried that my family would not have enough money to pay for necessities like clothing or bills. | | | | **0.45** | |
| CHAOS - We were usually able to stay on top of things. | | | | **0.47** | |
| QUIC - At least one of my parents was disorganized. | | | | **0.48** | |
| CHAOS - We could usually find things when we needed them. | | | | **0.54** | |
| CHAOS - No matter how hard we tried, we always seemed to be running late. | | | | **0.56** | |
| QUIC - I lived in a clean house. | | | | **0.64** | |
| QUIC - In my house things I needed were often misplaced so that I could not find them. | | | | **0.69** | |
| QUIC - I lived in a cluttered house (e.g., piles of stuff everywhere). | | | | **0.75** | |
| QUIC - One of my parents was unemployed and couldn't find a job even though he/she wanted one. | | | | | **0.32** |
| QUIC - I often wondered whether or not one of my parents would come home at the end of the day. | | | | | **0.34** |
| QUIC - At least one of my parents changed jobs. | | | | | **0.38** |
| QUIC - There were people coming and going in my house that I did not expect to be there. | | | | | **0.41** |
| Perceived - People often moved in and out of my house on a pretty random basis. | | | | | **0.48** |
| Romantic partners - father | | | | | **0.53** |
| Romantic partners - mother | | | | | **0.58** |
| QUIC - I experienced changes in my custody arrangement. | | | | | **0.60** |
| QUIC - There were long periods of time when I didn't see one of my parents (e.g. military deployment, jail time, custody arrangements). | | | | | **0.62** |
| QUIC - At least one of my parents had many romantic partners. | | | | | **0.65** |
| Perceived - My parents had a difficult divorce or separation during this time. | | | | | 0.67 |
| CHAOS - The telephone took up a lot of our time at home. | | | | | |
| QUIC - My parents were very late to pick me up (e.g., from school, aftercare or sports). | | | | | |
| QUIC - At least one of my parents would plan something for the family, but then not follow through with the plan. | | | | | |
| QUIC - I worried that I was not going to have enough food to eat. | | | | | |

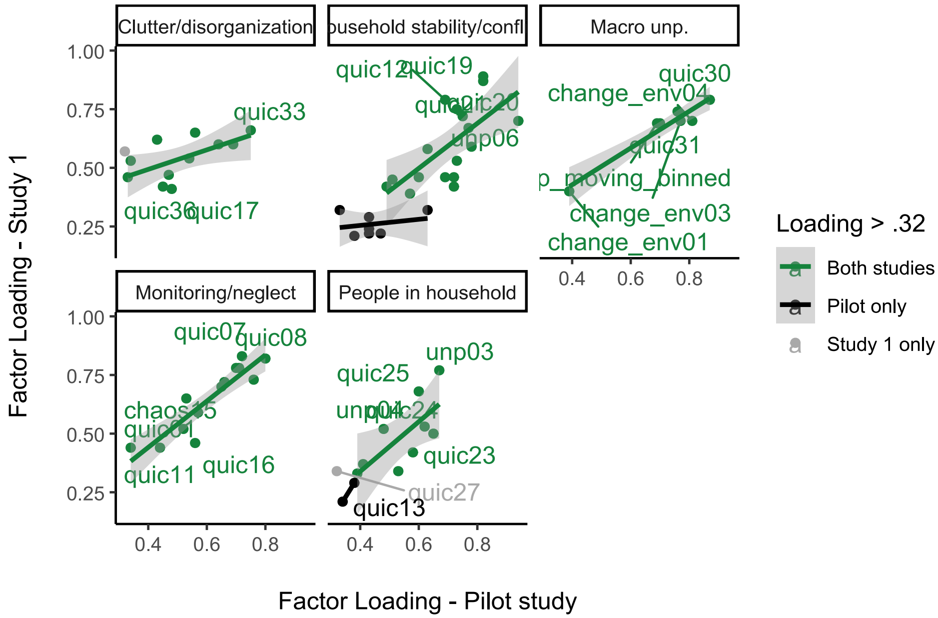



**Figure A3.1.** Comparison of factor loadings of unpredictability items in the Pilot and Study 1.

**Study 1.** Similar to the Pilot, the EFA yielded five factors based on parallel analysis. We plotted the factor loadings of each factor in the Pilot against those in Study 1 to investigate their correspondence (See Figure A3.1). In general, individual items largely loaded on the same factors, and the sizes of their loadings were also comparable. The items from the CHAOS were found to be most unstable, with many showing a loading < .32 in one of the two studies.

### Bivariate correlations between future orientation and impulsivity with attention tasks

Table A3.2 shows bivariate correlations between self-reported depression, impulsivity, future orientation, and SES with each of the Flanker SSP parameters and the Global-Local drift rate difference. Participants who reported more depressive symptoms had a lower strength of perceptual input. Participants who reported more impulsivity had a lower strength of perceptual input, higher interference, as well as a more holistic processing style. Participants who were more future oriented had a higher strength of perceptual input and a more detail-oriented processing style, without an association with interference.

**Table A3.3.** Bivariate correlations between exploratory measures and SSP parameters.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Depression | - | 1527 | 1527 | 545 | 545 | 1527 | 1527 | 1527 | 1527 | 445 |
| 2. Impulsivity | 0.37*** | - | 1527 | 545 | 545 | 1527 | 1527 | 1527 | 1527 | 445 |
| 3. Future orientation | -0.17** | -0.51** | - | 545 | 545 | 1527 | 1527 | 1527 | 1527 | 445 |
| 4. Internal attention style | 0.49*** | 0.46*** | -0.18** | - | 545 | 545 | 545 | 545 | 545 | 445 |
| 5. External attention style | 0.40*** | 0.30*** | -0.12** | 0.69*** | - | 545 | 545 | 545 | 545 | 445 |
| 6. Flanker - Perceptual input | -0.06* | -0.07** | 0.10*** | -0.04 | -0.02 | - | 1527 | 1527 | 1527 | 445 |
| 7. Flanker - Interference | 0.04 | 0.07** | -0.04 | 0.07 | 0.07 | -0.09** | - | 1527 | 1527 | 445 |
| 8. Flanker - boundary separation | -0.00 | 0.01 | 0.03 | 0.01 | 0.02 | 0.39*** | -0.05 | - | 1527 | 445 |
| 9. Flanker - Non-decision time | 0.03 | 0.02 | -0.02 | -0.05 | -0.05 | -0.26** | -0.03 | -0.50** | - | 445 |
| 10. Global-Local - Drift rate difference | 0.01 | 0.11* | -0.12* | 0.07 | 0.06 | -0.13** | -0.01 | 0.01 | 0.03 | - |

*Note:* * = $p < .05$; ** = $p < .01$; *** = $p < .001$. The upper diagonal presents sample sizes for each comparison.

## Section 3. Model fit

### Pilot

**Cueing and Change Detection Task.** In our initial, preregistered approach, DDM models for the Cueing and Change Detection Task were fit with the fast-dm-30 software (Voss et al., 2015) using maximum likelihood (ML) estimation. For both tasks, we started out with a model that freely estimated all parameters, and then fit additional models with an increasing number of constrained parameters. We compared model fit using the Bayesian information criterion (BIC), for which smaller values indicate better fit. For the Change Detection Task, the most simple model provided the best fit. This model freely estimated the drift rate, non-decision time, and boundary separation, and fixed all other parameters.

For the Cued Attention Task, three models provided comparable model fit. However, all three models showed estimation problems, especially with regard to the boundary separation. Specifically, boundary separation estimates for several participants ended up at an upper boundary of 10, indicating that they were not recovered well (see Figure A3.2). Based on subsequent external input, we fit an additional model using Kolmogorov-Smirnov (KS) estimation instead of ML estimation, additionally estimating the inter-trial variability parameter of the non-decision time. This improved parameter estimation (see Figure A3.4).

Finally, we switched to estimation using Hierarchical Bayesian DDM (HDDM) for our final analyses. The main reason for this step was that although KS estimation seemed to work well, we had fewer trials than is typically recommended for this estimation technique (Lerche et al., 2017). An advantage of HDDM is that it uses group-level estimates to inform and constrain individual-level estimates. This is especially useful in cases such as ours, where we have a large sample size but relatively few trials per participant.

The HDDM models were fit using the *runjags* package (Denwood, 2016), using code from D. J. Johnson et al. (2017). All models were fit using three Markov Chain Monte Carlo (MCMC) chains. Each of these chains started with 2,000 burn-in samples,

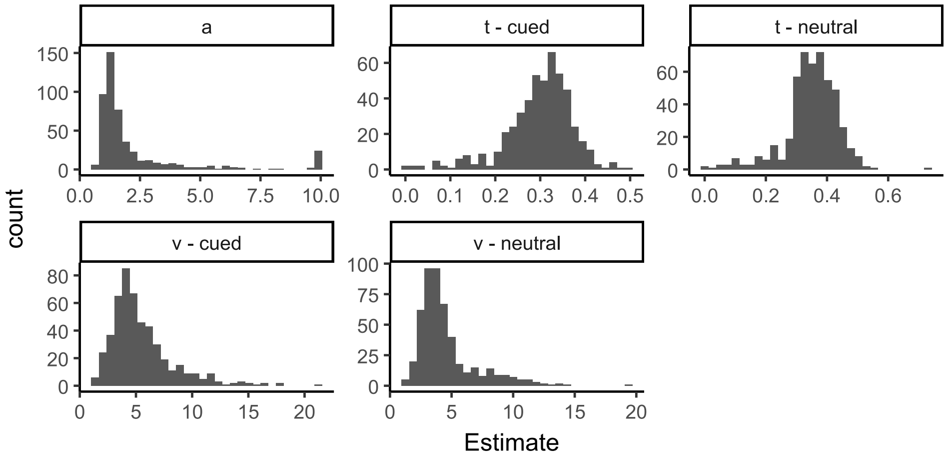



**Figure A3.2.** Distributions of Drift Diffusion Model parameters for the Cued Attention Task using Maximum likelihood estimation.

followed by 10,000 additional samples. To decrease the total size of the model, every 10th sample was retained, resulting in a posterior sample of 3,000 samples.

Model convergence was assessed (1) by visually inspecting the traces, which should not contain any drifts or large jumps (see Figure A3.4 and Figure A3.7); (2) through simulation. Specifically, we used each participant's DDM estimates to simulate

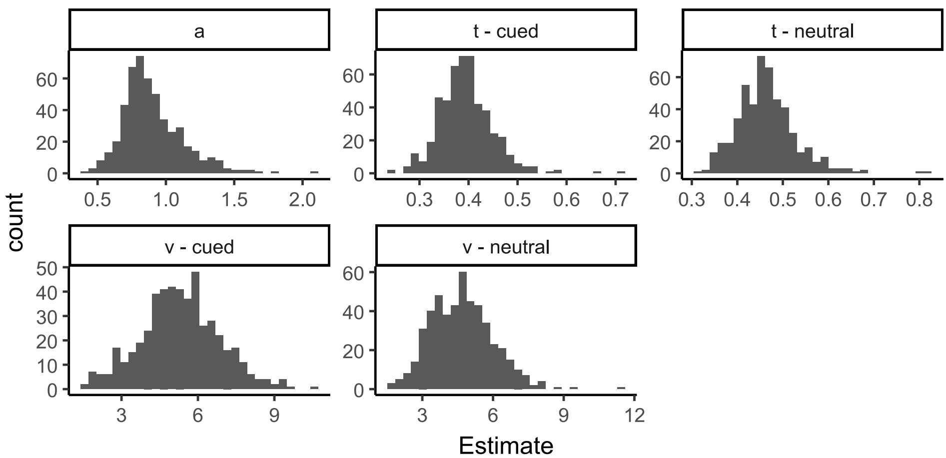



**Figure A3.3.** Distributions of Drift Diffusion Model parameters for the Cued Attention Task using Kolmogorov-Smirnov estimation.

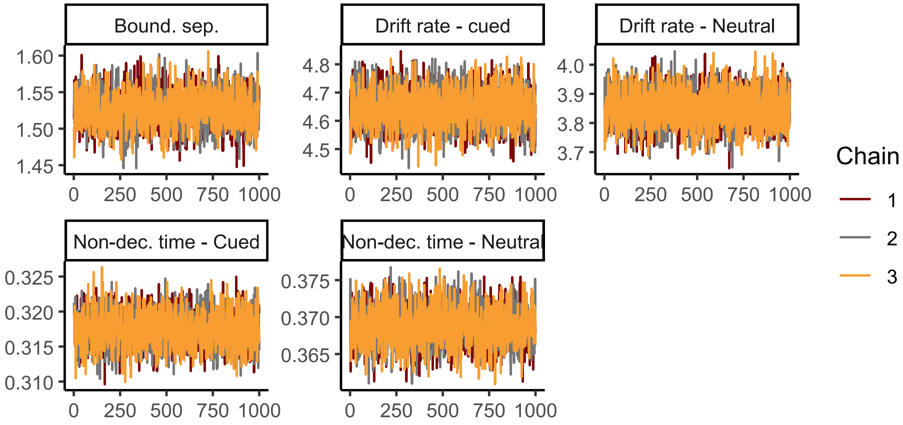



**Figure A3.4.** Chain convergence for the Hierarchical Bayesion Drift Diffusion Model for the Cued Attention Task. Plots should resemble a 'fat, hairy catterpillar'.

100 RT and accuracy estimates (per condition). The distributions of the participant's true RTs and their simulated RTs were assessed through bivariate correlations at the 25th, 50th and 75th percentile (See Figure A3.5 and Figure A3.8). We made the same comparison for mean accuracy levels (See Figure A3.6 and Figure A3.9).

**Flanker.** To fit the SSP model to the Flanker data, we followed recommendations by Grange (2016). First, we searched for the optimal set of starting values. For each participant, we used 50 sets of starting parameters with a variance of 20 for each, simulating 1,000 trials. After finding the optimal starting values, we fit the final model based on 50,000 simulated trials. Model fit was assessed through simulation. For each participant, we simulated 50,000 trials. We then calculated correlations between observed and simulated RTs at the 25th, 50th, and 75th percentile, as well as between observed and simulated mean accuracy. As can be seen in Figure A3.10 and Figure A3.11, we observed high agreement between observed and simulated RTs and accuracy rates.

**Study 1**

Model fit of the Flanker was done the same as in the Pilot. Figure A3.12 and Figure A3.13 show the model fit based on simulated data. We found good model fit across all three conditions, both for RTs as well as accuracy rates.

**Study 2.**

**Flanker**. Model fit of the Flanker was done the same as in the Pilot and Study 1. Figure A3.14 and Figure A3.15 show the model fit based on simulated data. We found good model fit, both for RTs as well as accuracy rates.

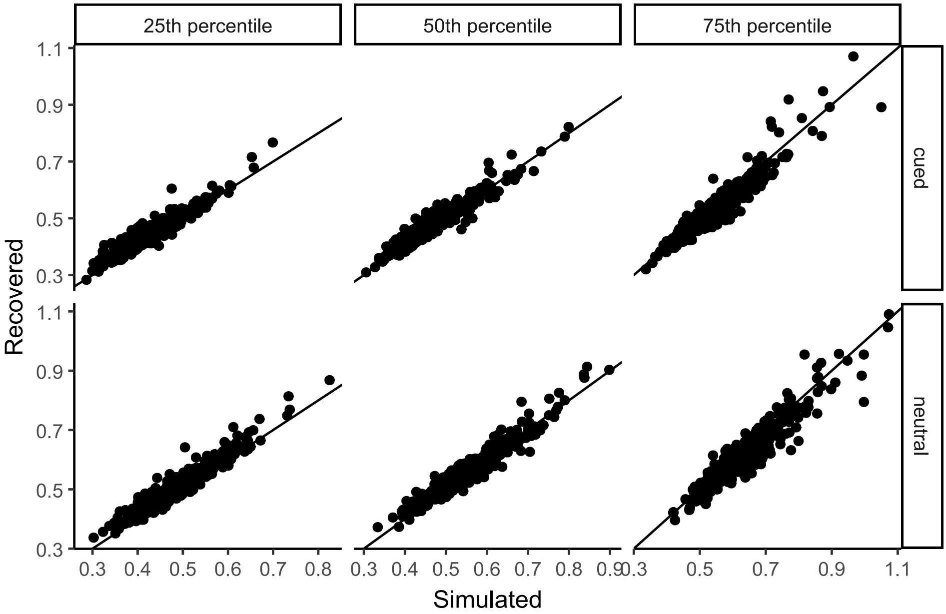



**Figure A3.5.** Drift Diffusion parameter recovery for the Cued Attention Task comparing simulated and recovered response times at the 25th, 50th and 75th percentile of response times.

## Section 4. Deviations from preregistrations

In this section, we provide a numbered overview of the deviations from the preregistration in each study.

### Pilot

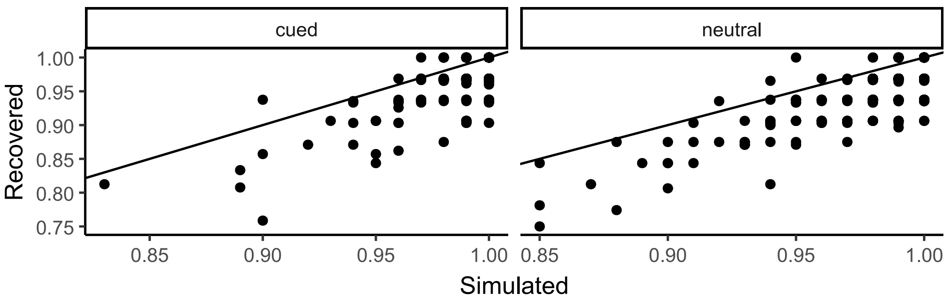



**Figure A3.6.** Drift Diffusion parameter recovery for the Cued Attention Task comparing simulated and recovered accuracy rates.

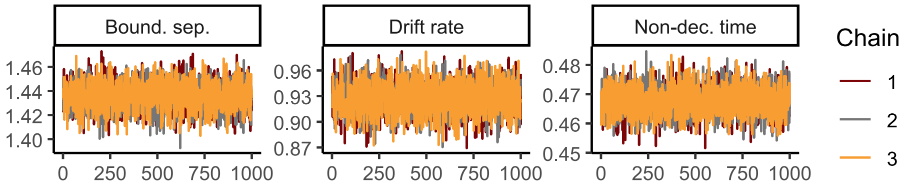



**Figure A3.7.** Chain convergence for the Hierarchical Bayesion Drift Diffusion Model for the Change Detection Task. Plots should resemble a 'fat, hairy catterpillar'.

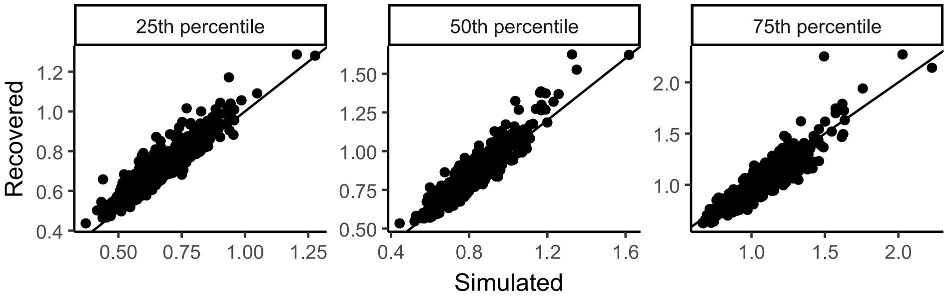



**Figure A3.8.** Drift Diffusion parameter recovery for the Change Detection Task comparing simulated and recovered response times at the 25th, 50th and 75th percentile of response times.

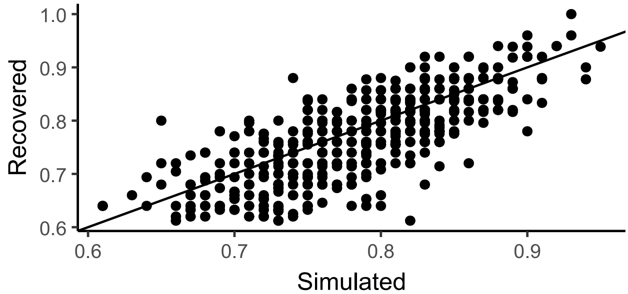



**Figure A3.9.** Drift Diffusion parameter recovery for the Change Detection Task comparing simulated and recovered accuracy rates.

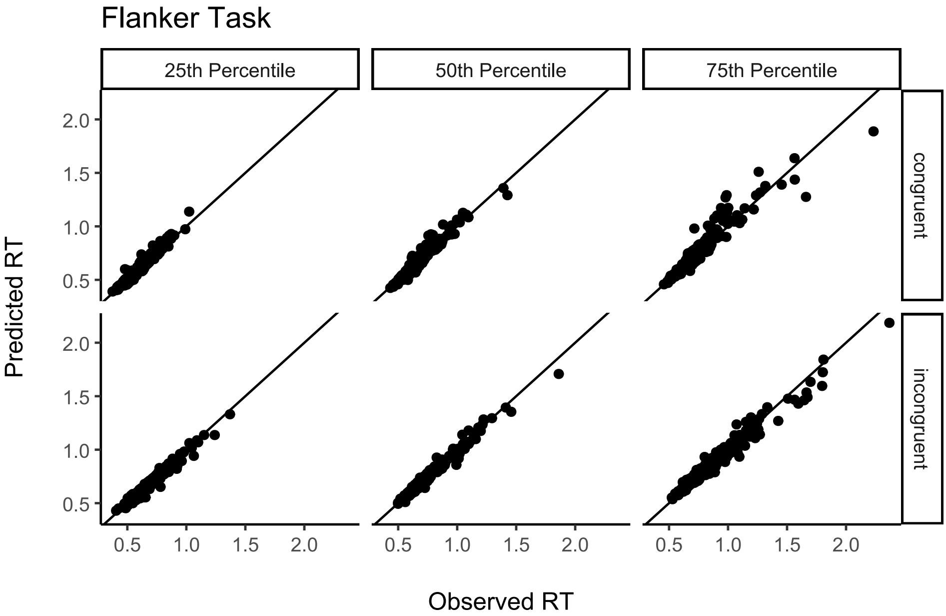



**Figure A3.10.** Drift Diffusion parameter recovery for the Flanker comparing simulated and recovered response times at the 25th, 50th and 75th percentile of response times.

1. *DDM estimation using Hierarchical Bayesian DDM instead of Maximum Likelihood estimation.* The rational for this change is explained in more detail in section 3.

2. *Focus on Flanker Interference instead of separate attention parameters.* The SSP model provides two parameters representing attentional processes: (1) the initial attention width, and (2) the rate at which attention shrinks towards the central target. We had initially planned to analyze both parameters separately. However, after analyzing the data from Study 1, we realized that the estimates of these two parameters separately were very unstable. We noticed this when plotting the within-person estimates between conditions (standard, enhanced, and degraded) against each other. Figure A3.16 provides an overview of these correlations for attentional width, shrinking rate, interference, and, for illustrative purposes, the RT difference score. We consider the comparison between the standard and enhanced condition the most informative, as the stimuli in these conditions were most similar. The within-person correlations between conditions of attentional width and shrinking were very low. However, the correlations were substantially higher for interference (which even outperformed the standard RT difference scores, as typically used in traditional assessments). Thus, we decided to use the Interference estimate in our analyses across all studies.

**Study 1**

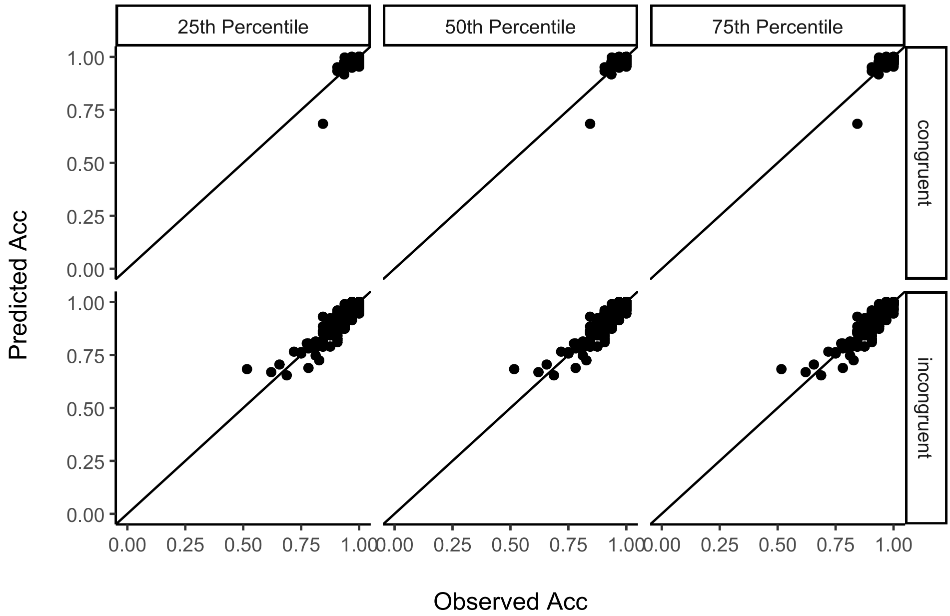



**Figure A3.11.** Drift Diffusion parameter recovery for the Flanker comparing simulated and recovered accuracy rates.

1. In our preregistration of Study 1, we preregistered four exploratory (hypothesis-generating) aims, which were unrelated to the primary (hypothesis-driven) aims described in the main manuscript. These were: (1) Investigating the factor structure of unpredictability measures and comparing it to the structure found in the Pilot; (2) Exploring the role of state anxiety, hunger, and sleep deprivation as potential moderators of the relationship between adversity and attention performance; (3) Exploring bivariate correlations between measures of adversity, attention, and measures of temporal orientation; (4) Explore the correlation between current depressive symptoms and retrospective measures of adversity. The results of aim 1 are described in Section 2. Results of aim 2-4 are described in Section 2.

**Study 2**

1. *Global-Local performance.* In the original preregistration, we specified that we would exclude participants who performed at chance on either the Flanker Task or the Global-Local Task, which was defined as an accuracy of 59.4% or lower. However, initial inspections of the Global-Local Task data showed that a substantial part of the sample did not reach this cut-off, suggesting that the task was more difficult than anticipated. Thus, we developed a more fine-grained approach (described below) in an attempt to distinguish between 1) participants who did not understand the task and 2) participants who understood the task, but found it difficult to perform well. Given the assumptions of DDM, the first group would have to be excluded because

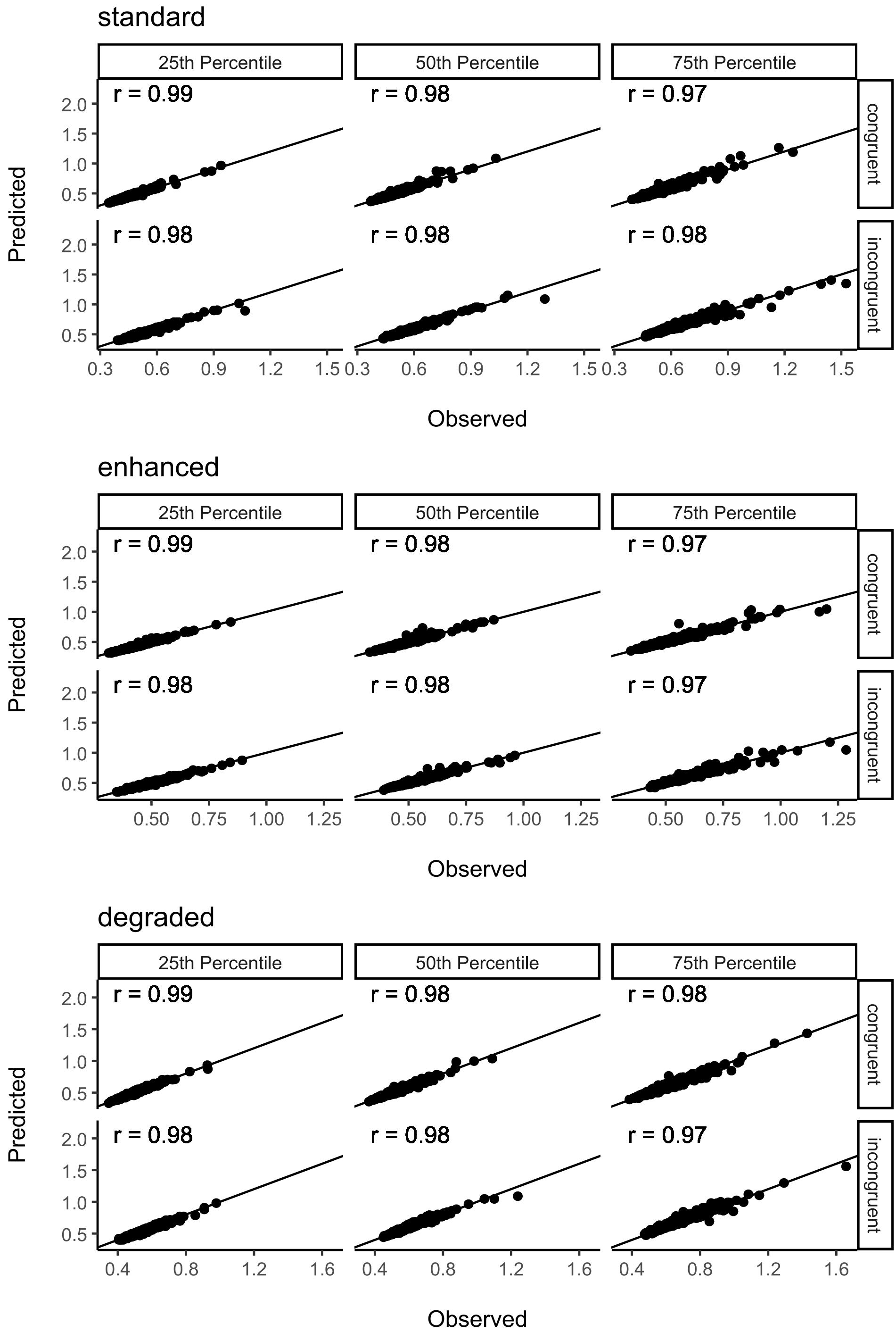



**Figure A3.12.** Drift Diffusion parameter recovery for the Flanker comparing simulated and recovered response times at the 25th, 50th and 75th percentile of response times across the standard, enhanced, and degraded condition.

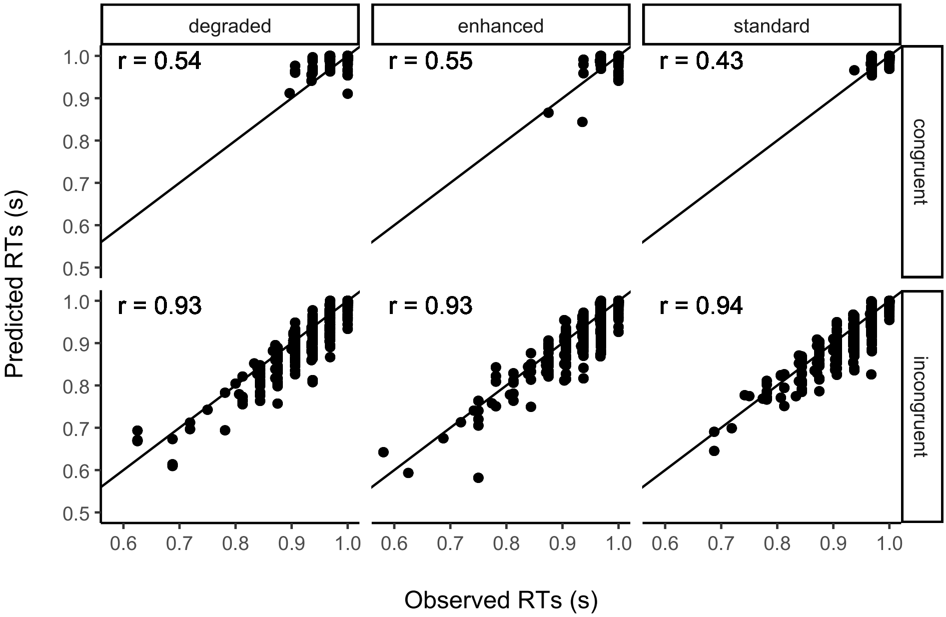



**Figure A3.13.** Drift Diffusion parameter recovery for the Flanker comparing simulated and recovered accuracy rates across the standard, enhanced, and degraded condition.

they likely did not go through a process of information accumulation. However, the model should be able to adequately fit the data of the second group.

The amended analysis approach for the Global-Local Task looked as follows: (1) Fit the data to the cleaned data of the full sample, including participants who performed at or below chance level (i.e., after trial-level exclusions but before case-wise exclusions); (2) Based on recovered parameter estimates for each participant, simulate the same number of trials (reaction times and accuracy) using the `RWiener` package, separately for Global and Local trials. (3) For each participant, calculate the 25th, 50th and 75th quantile of both their real RTs and the simulated RTs. In addition, we calculate mean accuracies based on the real and simulated data. (4) Compute standardized residuals between the real and simulated data for RTs at each quantile and for accuracy. In case of good fit, the residual should be close to zero. (5) Exclude the data of participants with any standardized residual > 3.2 SD.

2. *Multiverse analysis.* In the preregistration, we planned to include three variables as covariates that were previously featured as arbitrary exclusion decisions in the multiverse specification: 1) whether or not participants rescaled the screen; 2) whether or not participants exited fullscreen mode at some point during the tasks; 3) Whether or not participants experienced interruptions during the tasks. Our rea-

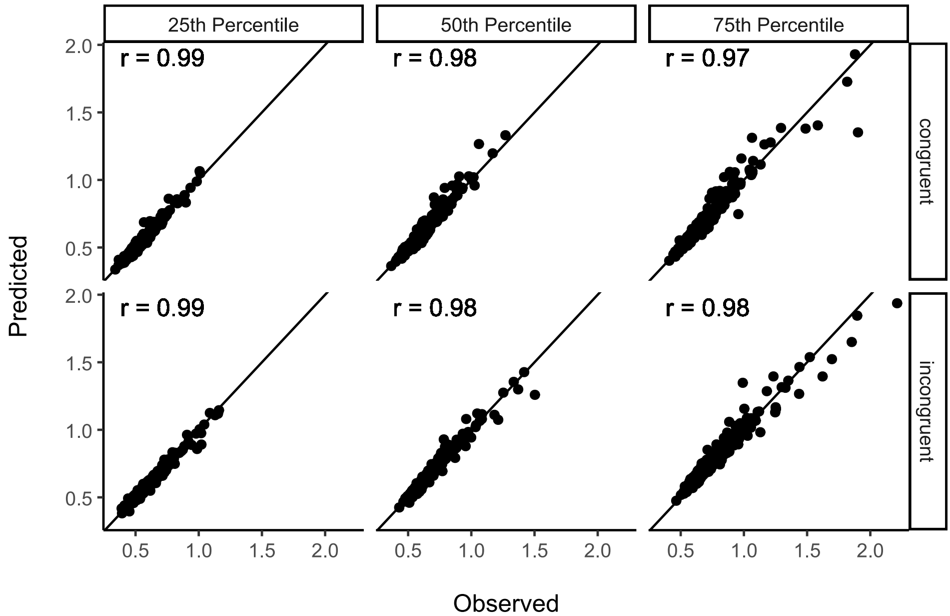



**Figure A3.14.** Drift Diffusion parameter recovery for the Flanker comparing simulated and recovered response times at the 25th, 50th and 75th percentile of response times.

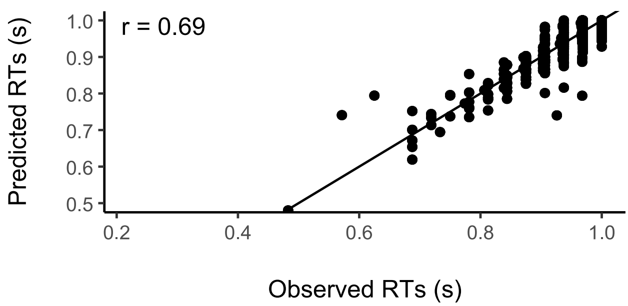



**Figure A3.15.** Drift Diffusion parameter recovery for the Flanker comparing simulated and recovered accuracy rates.

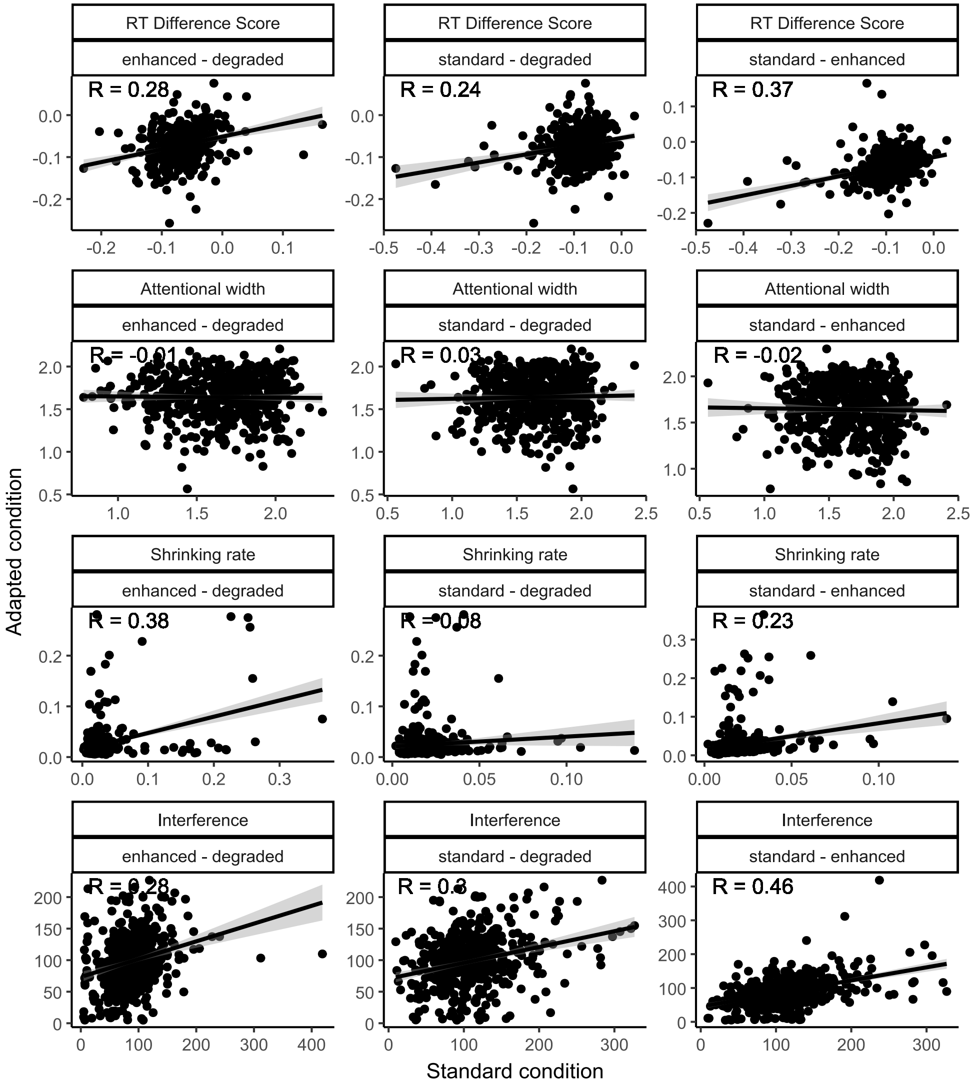



**Figure A3.16.** Within-person bivarate correlations between Flanker conditions in Study 2 for RT differences, attentional width, shrinking rate, and interference.

soning was that these three factors were consistently found to have a large impact on model results. However, we realized later on that adding these factors as covariates was not a good approach from a causal inference standpoint. That is, it is more likely that each of these factors added random noise to our estimates than that they had a causal effect on the outcome. Therefore, we decided instead to include these factors as arbitrary decisions in the multiverse analyses, similarly as the previous two studies. This allowed for a coherent assessment of influential factors across all three experiments.

**Section 5. Multiverse analysis.**

In this section, we provide additional results of the multiverse analyses. Specifically, we report (1) the distributions of *p*-values across the multiverses and (2) influential data cleaning decisions.

**Pilot**

Figure A3.17 and Figure A3.18 present *p*-distributions and the explained variance of each data cleaning decision in the variation in effect sizes for the results presented in Table 3 in the main text.

**Study 1**

Figure A3.19 and Figure A3.20 present *p*-distributions and the explained variance of each data cleaning decision in the variation in effect sizes for the results presented in Table 4 in the main text.

Figures A3.21-A24 present *p*-distributions and the explained variance of each data cleaning decision in the variation in effect sizes for the results presented in Table 5 in the main text.

**Study 2**

Figure A3.25 and Figure A3.26 present *p*-distributions and the explained variance of each data cleaning decision in the variation in effect sizes for the results presented in Figure 2 in the main text.

Figure A3.27 and Figure A3.28 present *p*-distributions and the explained variance of each data cleaning decision in the variation in effect sizes for the results presented in Figure 3 in the main text.

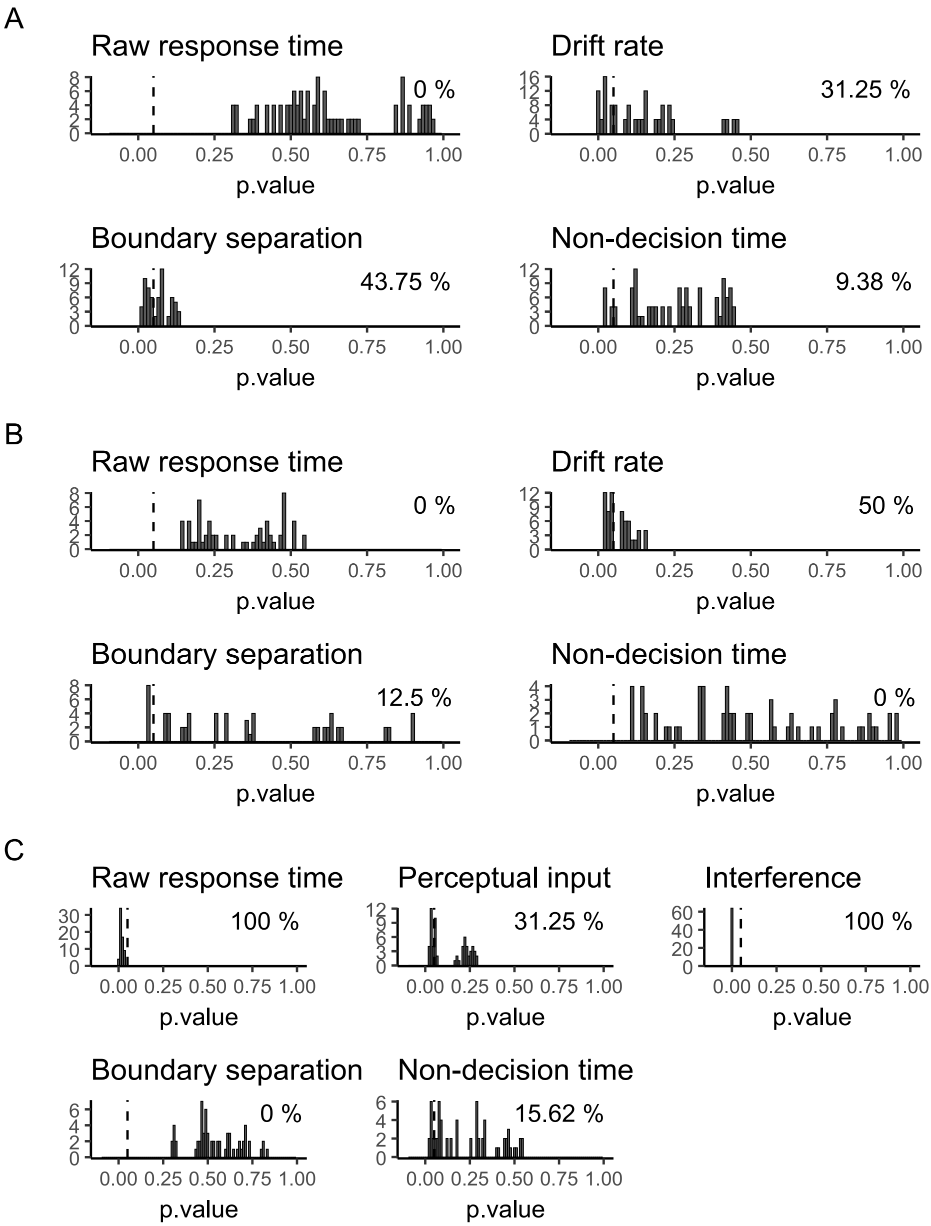



**Figure A3.17.** Multiverse p-value distributions belonging to the analyses reported in Table 3 in the main text. The dashed vertical line depicts the cut-off of .05. The percentages in the upper-right corners are the percentage of statistically significant analyses in multiverse. Panel A: Cued Attention Task. Panel B: Change Detection Task. Panel C: Flanker Task.

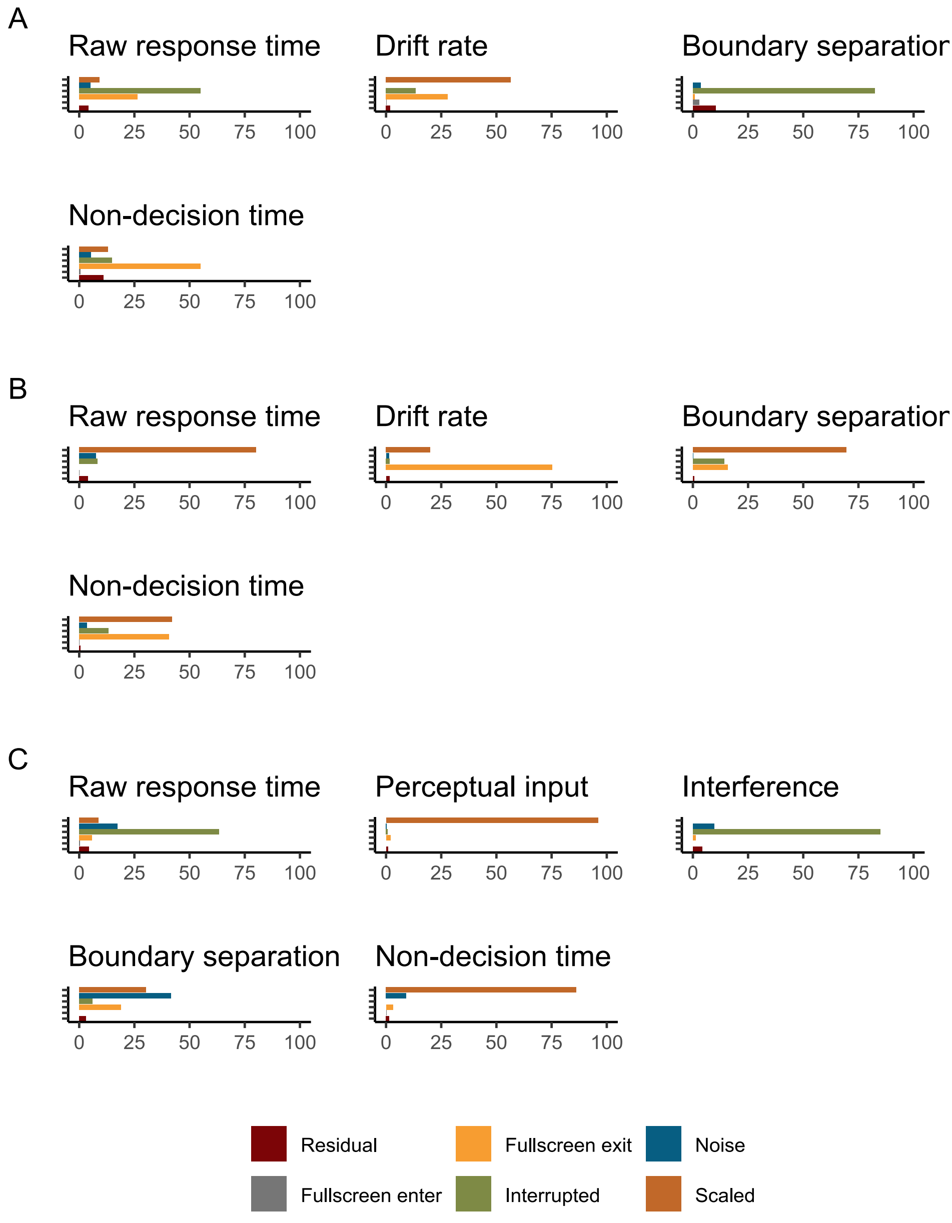



**Figure A3.18.** Multiverse explained variance of each data cleaning decision belonging to the analyses reported in Table 3 in the main text. Panel A: Cued Attention Task. Panel B: Change Detection Task. Panel C: Flanker Task.

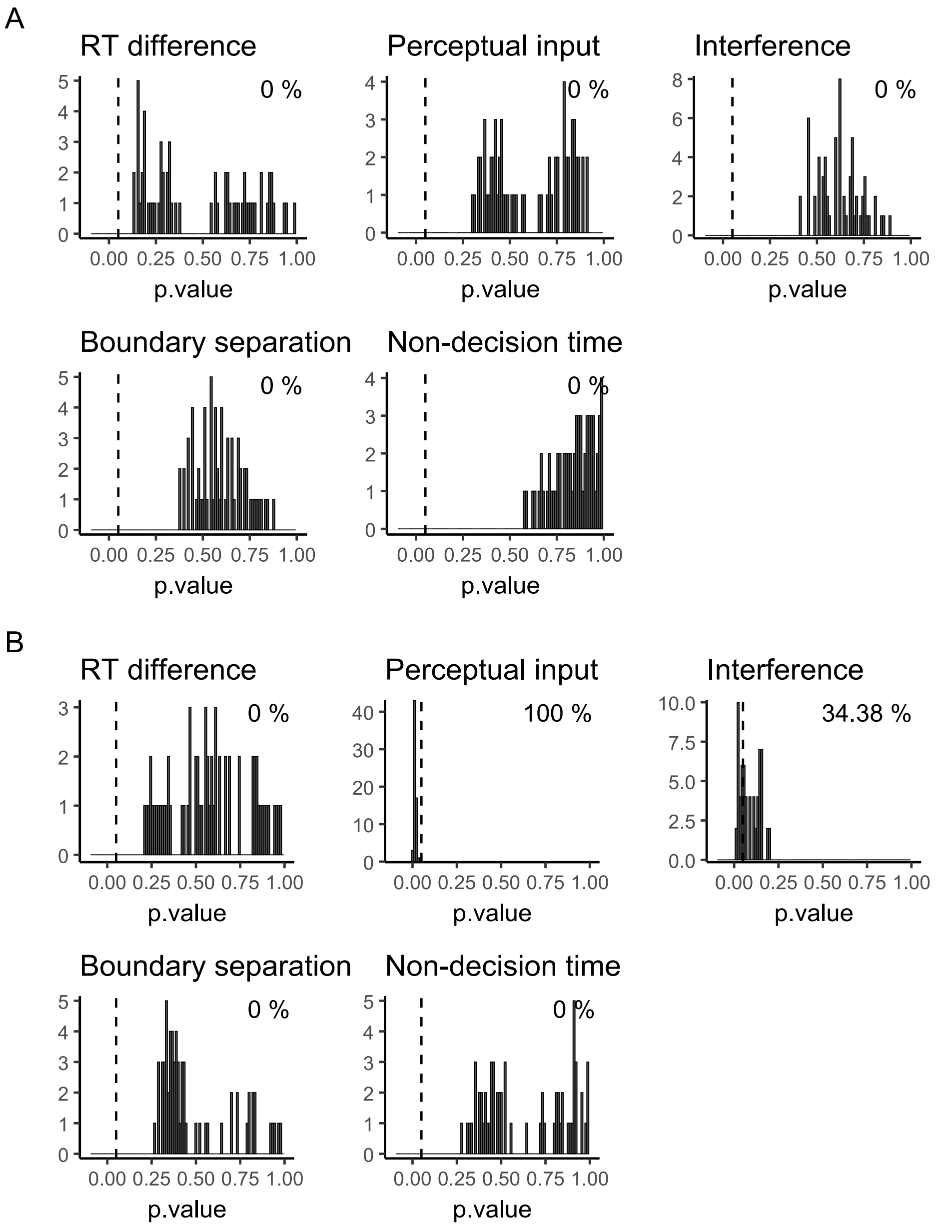



**Figure A3.19.** Multiverse p-value distributions belonging to the analyses reported in Table 4 in the main text. The dashed vertical line depicts the cut-off of .05. The percentages in the upper-right corners are the percentage of statistically significant analyses in multiverse. Panel A: Analyses involving violence exposure (hypothesis-driven). Panel B: Analyses involving unpredictability (exploratory).

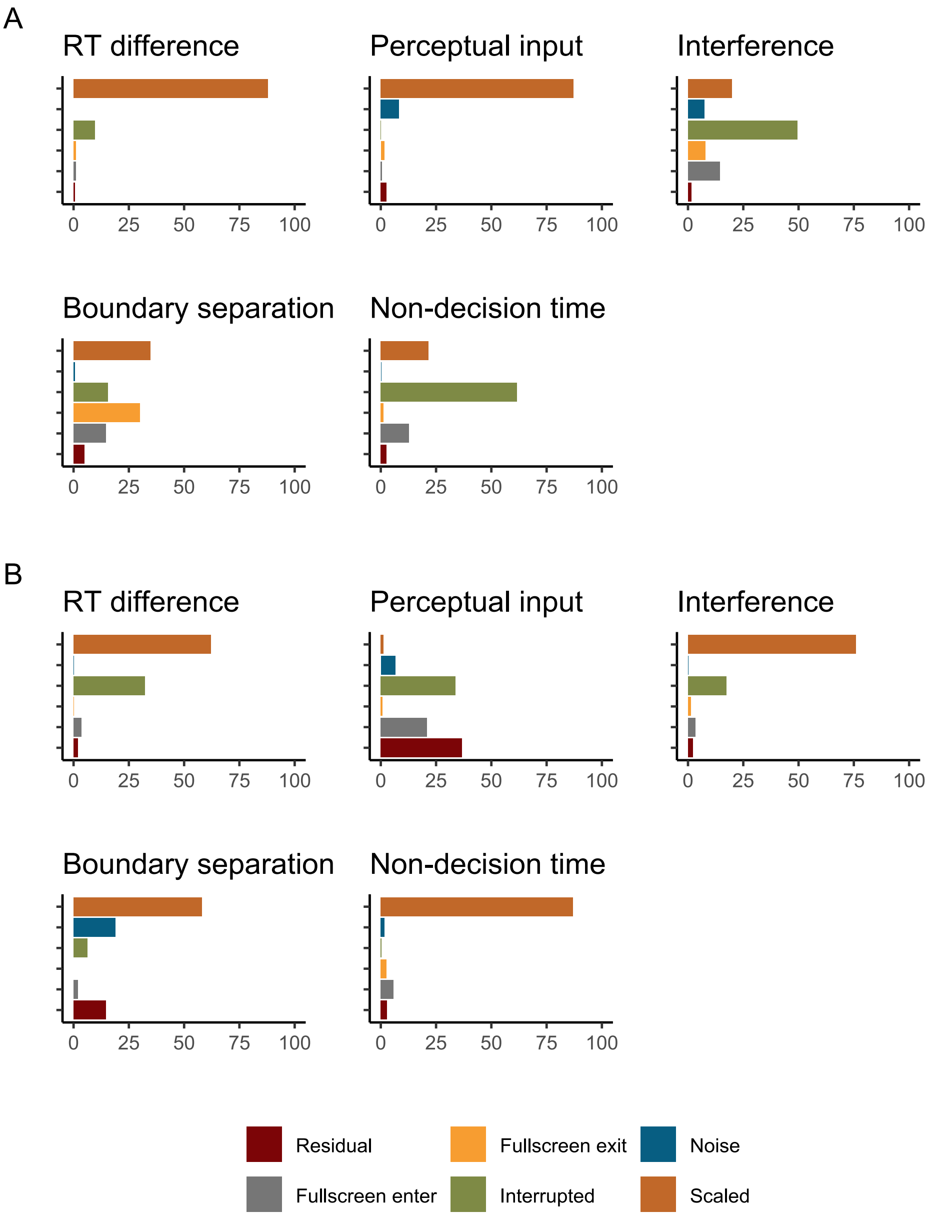



**Figure A3.20.** Multiverse explained variance of each data cleaning decision belonging to the analyses reported in Table 4 in the main text. Panel A: Analyses involving violence exposure (hypothesis-driven). Panel B: Analyses involving unpredictability (exploratory).

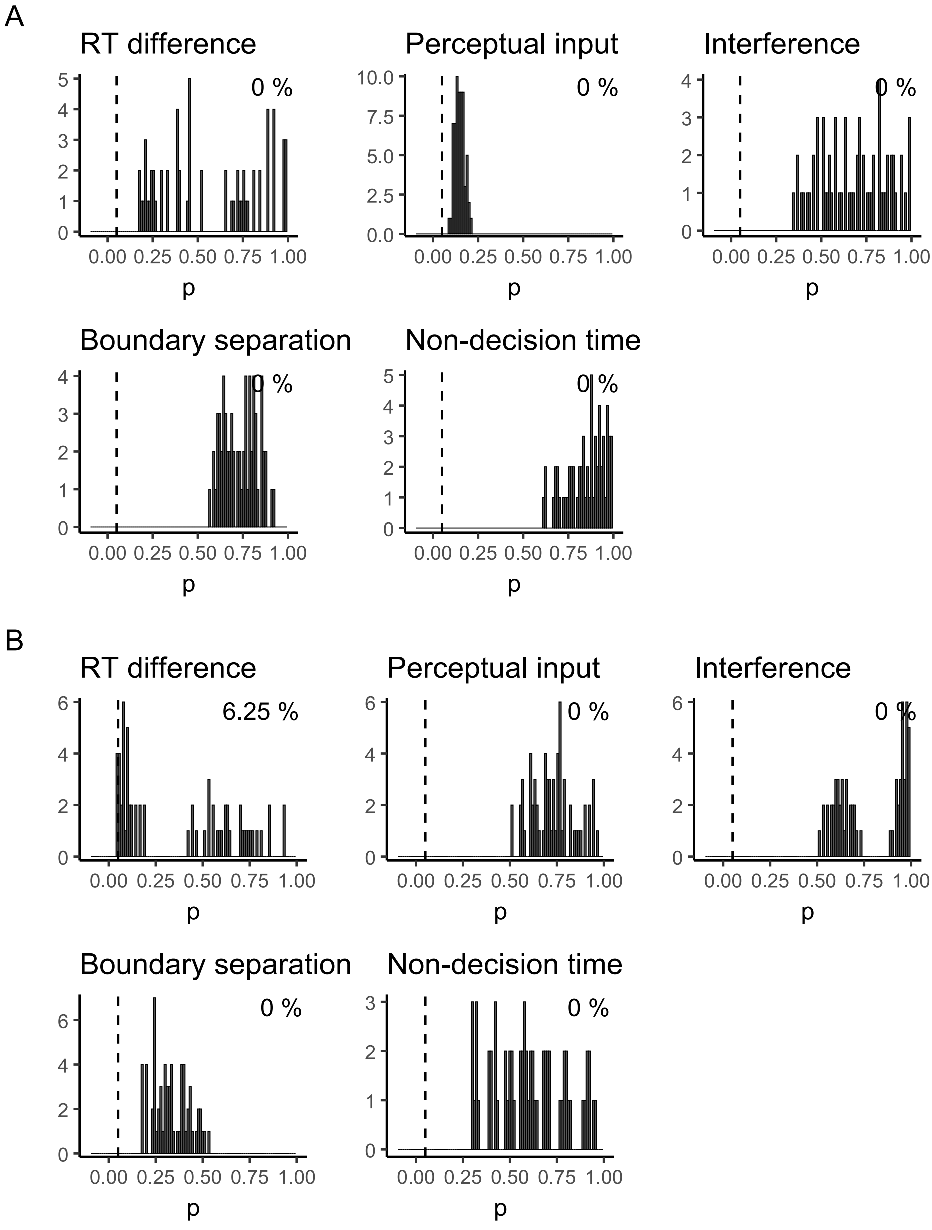



**Figure A3.21.** Multiverse p-value distributions belonging to the interaction effects involving violence exposure reported in Table 5 in the main text. The dashed vertical line depicts the cut-off of .05. The percentages in the upper-right corners are the percentage of statistically significant analyses in multiverse. Panel A: Analyses comparing the enhanced condition to the standard condition. Panel B: Analyses comparing the degraded condition to the standard condition.

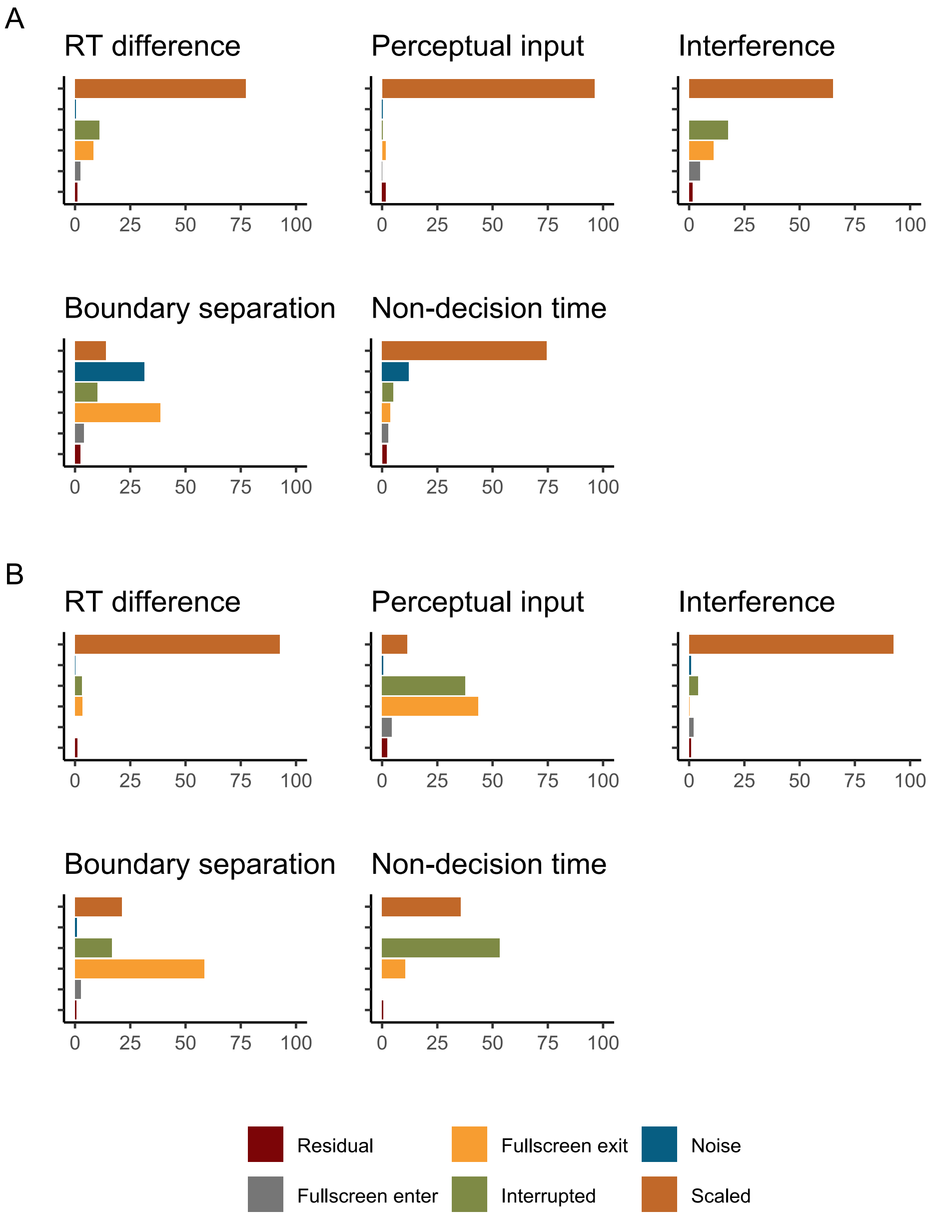



**Figure A3.22.** Multiverse explained variance of each data cleaning decision belonging to the interaction effects involving violence exposure reported in Table 5 in the main text. Panel A: Analyses comparing the enhanced condition to the standard condition. Panel B: Analyses comparing the degraded condition to the standard condition.

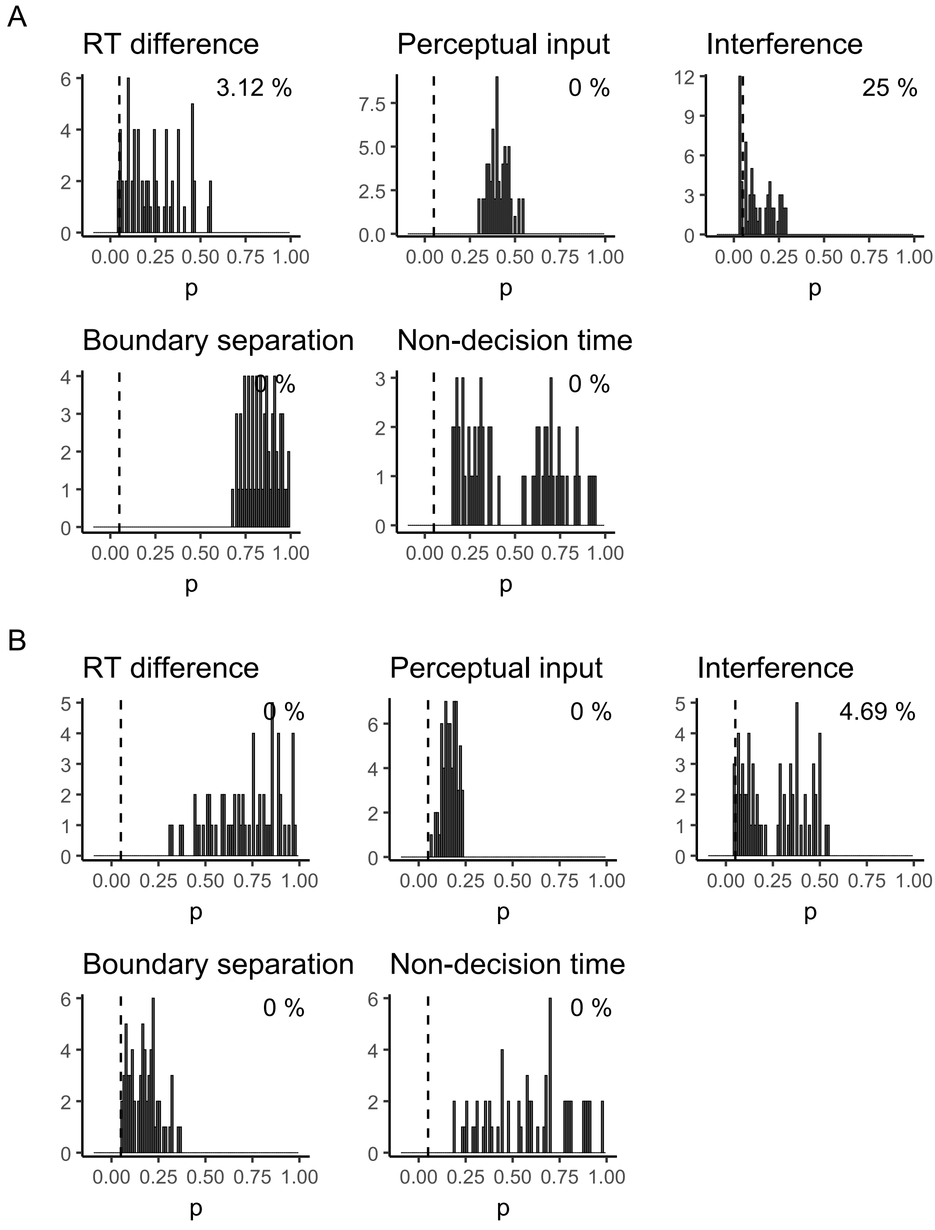



**Figure A3.23.** Multiverse p-value distributions belonging to the interaction effects involving unpredictability reported in Table 5 in the main text. The dashed vertical line depicts the cut-off of .05. The percentages in the upper-right corners are the percentage of statistically significant analyses in multiverse. Panel A: Analyses comparing the enhanced condition to the standard condition. Panel B: Analyses comparing the degraded condition to the standard condition.

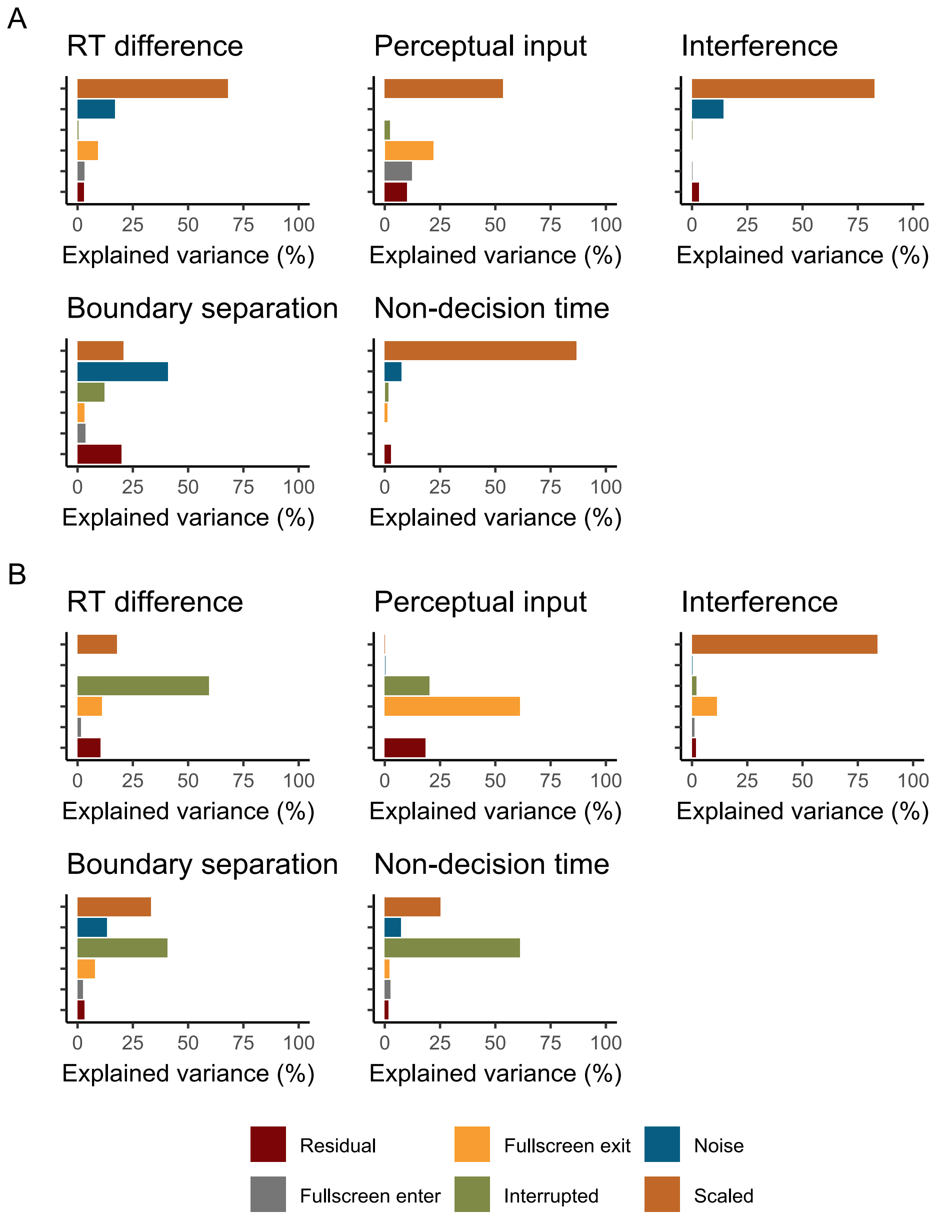



**Figure A3.24.** Multiverse explained variance of each data cleaning decision belonging to the interaction effects involving unpredictability reported in Table 5 in the main text. Panel A: Analyses comparing the enhanced condition to the standard condition. Panel B: Analyses comparing the degraded condition to the standard condition.

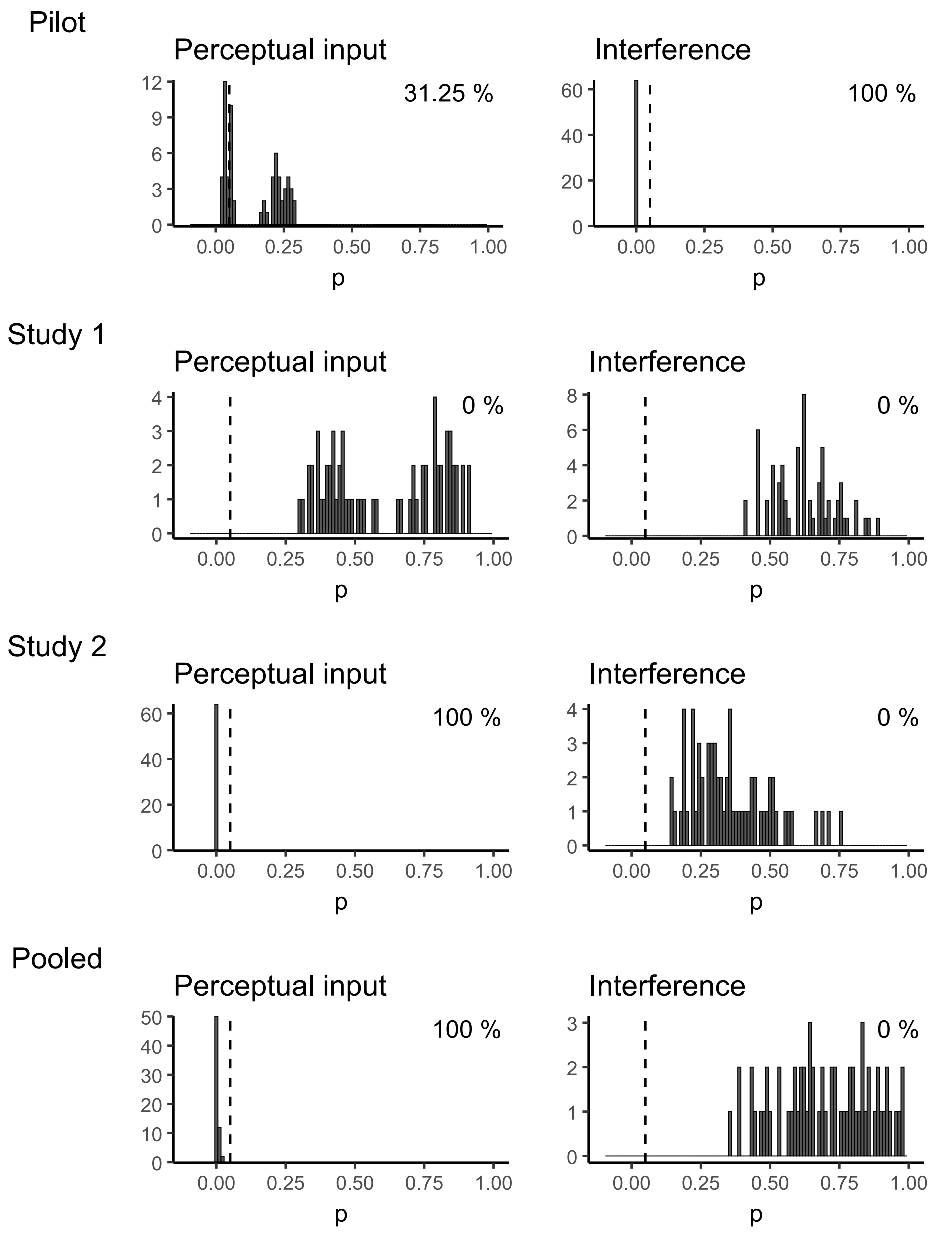



**Figure A3.25.** Multiverse p-value distributions belonging to the associations between violence exposure with perceptual input and interference in the Pilot, Study 1, Study 2, and pooled across all studies, as reported in Figure 2 in the main text.

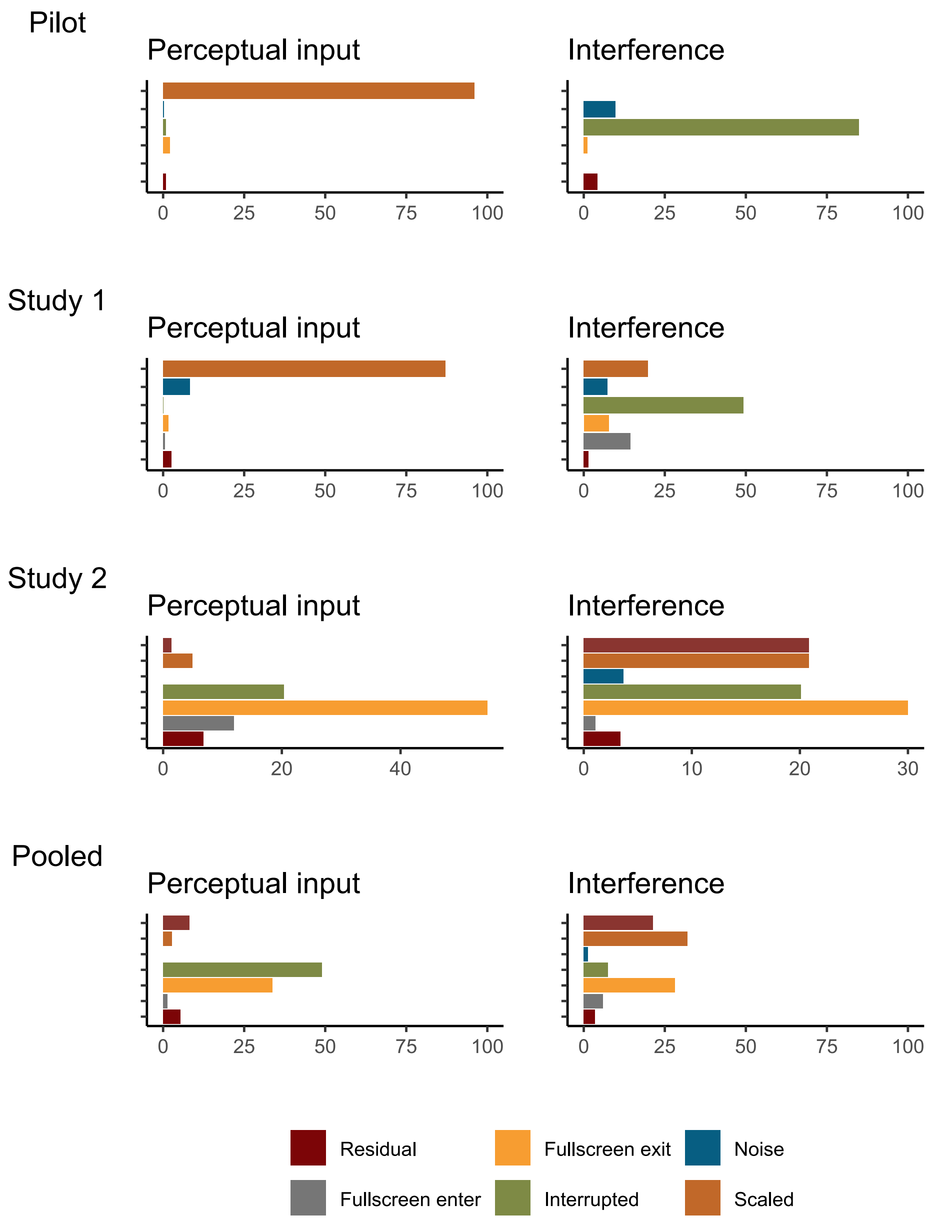



**Figure A3.26.** Multiverse explained variance of each data cleaning decision belonging to the associations between violence exposure with perceptual input and interference in the Pilot, Study 1, Study 2, and pooled across all studies, as reported in Figure 2 in the main text.

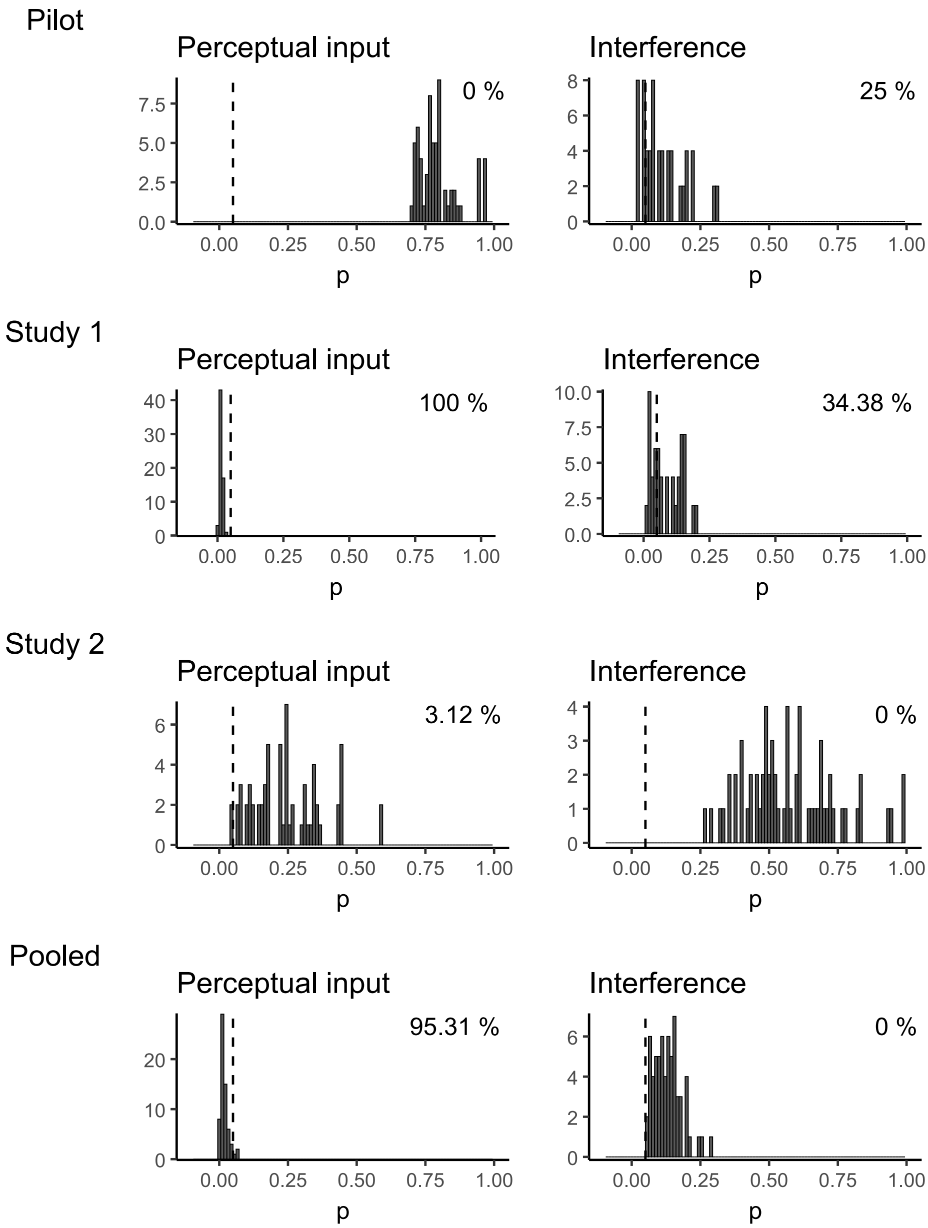



**Figure A3.27.** Multiverse p-value distributions belonging to the associations between unpredictability with perceptual input and interference in the Pilot, Study 1, Study 2, and pooled across all studies, as reported in Figure 3 in the main text.

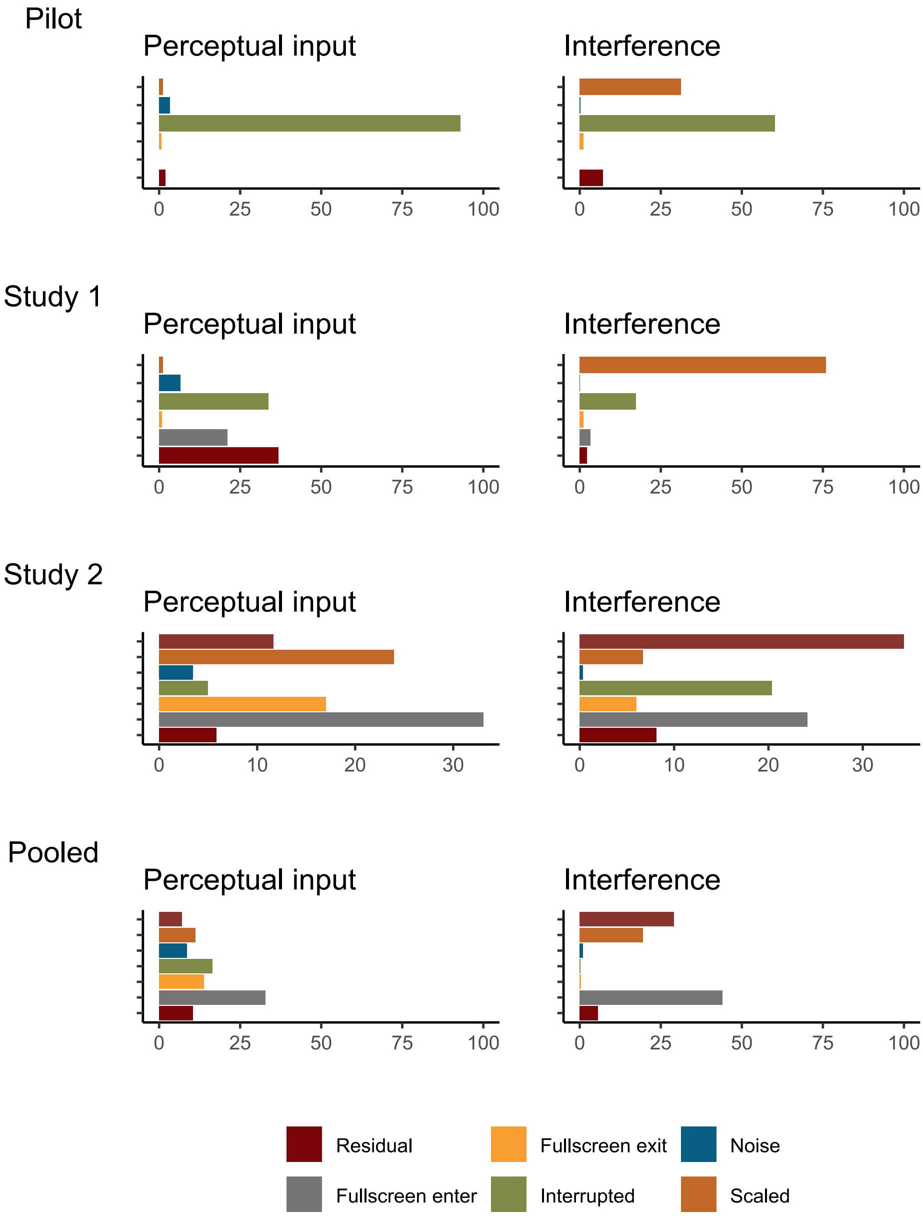



**Figure A3.28.** Multiverse explained variance of each data cleaning decision belonging to the associations between unpredictability with perceptual input and interference in the Pilot, Study 1, Study 2, and pooled across all studies, as reported in Figure 3 in the main text.

# Appendix 4 - Chapter 5

## Section 1. Pilot study

A total of 99 Dutch adolescents (mean age = 29.4, SD = 6.9, range = [20, 53]) participated in the Pilot study via Prolific. The main goal of the Pilot study was to obtain feedback on the tasks (e.g., difficulty, clarity of instructions), and to explore bivariate correlations between the measures. Participants completed the same tasks as in the main study: (1) Operation Span Task, (2) Rotation Span Task, and (3) Binding-Updating Task. The Rotation Span Task was administered in a second session, and was completed by a subsample of 50 participants. All three tasks followed the exact same procedure as in the main study.

In addition, participants completed measures of material deprivation, neighborhood threat, and unpredictability. These measures differed from the more comprehensive measures used in the main study, and were included to obtain quick, descriptive estimates. Material deprivation was measured using seven items about perceived level of available resources. Unpredictability was measured using a scale of perceived unpredictability (Mittal et al., 2015; Young et al., 2018). Neighborhood violence exposure was measured using the Neighborhood Violence Scale (NVS; Frankenhuis, Vries, et al., 2020; Frankenhuis & Bijlstra, 2018) as well as two items measuring involvement in fights. Participants responded to items of all questionnaires on a scale of 1 (never true) to 5 (very often true). Finally, participants provided feedback on the difficulty of the tasks and the clarity of the task instructions.

Pilot data were collected sequentially to allow for intermediate changes to instructions based on participants' feedback. The first session (including the Operation Span Task and the Binding-Updating Task) took approximately 35 minutes to complete, and participants were paid 5.25 GBP. The second session (including the Rotation Span Task) took approximately 9 minutes to complete, and participants were paid 1.50 GBP.

Table A4.1 presents bivariate correlations among the WM tasks, and between the WM tasks and measures of adversity. The WM tasks correlated moderately to strongly with each other. The strongest correlation was between the Binding and Updating score (.80). This is not surprising given that both scores are derived from the same task, and shows the importance of accounting for this association in the model. Neither unpredictability nor material deprivation were significantly associated with performance on any of the WM tasks. However, higher levels of experienced neighborhood threat were associated with lower performance on the Binding and Updating Task. Note that these associations were based on raw task performance and not on latent estimates.

**Table A4.1.** Bivariate correlations between WM tasks and adversity measures in the Pilot

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| **WM tasks** | | | | | | | |
| 1. Operation Span Task | - | 45 | 96 | 96 | 96 | 96 | 96 |
| 2. Rotation Span Task | 0.38* | - | 45 | 45 | 45 | 45 | 45 |
| 3. Binding Task | 0.42*** | 0.41** | - | 96 | 96 | 96 | 96 |
| 4. Updating Task | 0.42*** | 0.51*** | 0.80*** | - | 96 | 96 | 96 |
| **Adversity** | | | | | | | |
| 5. Unpredictability | -0.14 | 0.19 | -0.16 | -0.04 | - | 96 | 96 |
| 6. Threat | -0.05 | -0.15 | -0.24* | -0.22* | 0.25* | - | 96 |
| 7. Material deprivation | 0.06 | 0.04 | 0.04 | -0.09 | -0.23* | -0.39*** | - |
| Mean | 0.82 | 0.76 | 0.84 | 0.75 | 2.18 | -0.00 | 3.79 |
| SD | 0.15 | 0.16 | 0.16 | 0.17 | 0.90 | 0.87 | 0.76 |
| Median | 0.85 | 0.76 | 0.89 | 0.78 | 1.94 | -0.33 | 4.00 |
| Min | 0.39 | 0.37 | 0.31 | 0.31 | 1.00 | -0.95 | 1.57 |
| Max | 1.00 | 0.98 | 1.00 | 1.00 | 5.00 | 3.54 | 5.00 |
| Skew | -0.97 | -0.71 | -1.34 | -0.78 | 0.83 | 1.51 | -0.88 |
| Kurtosis | 0.14 | -0.10 | 1.26 | -0.04 | 0.11 | 2.30 | 0.16 |

*Note:* The upper diagonal presents sample sizes for each bivariate comparison. The measures of unpredictability, threat, and material deprivation differ from those in the main study.

## Section 2. Exploratory analyses

To contextualize our confirmatory (preregistered) findings, we conducted three exploratory (non-preregistered) analyses. First, we explored associations between adversity and performance on the separate WM tasks using linear regression. Second, we constrained regression paths in the SEM to zero, as an alternative to the equivalence tests. Third, we computed Bayes Factors for the equivalence tests.

### Linear regression analyses

We estimated a total of five linear regression models, one per WM task. Each model included the same independent variables and covariates as the primary analysis. We adjusted for multiple testing across models involving the Rotation Span Task, Operation Span Task, and the binding trials of the Binding-Updating Task, and separately for the updating trials of the Binding-Updating Task, as the former three tasks are primarily conceptualized as WM capacity tasks. We also tested for practical equivalence in the same way as for the confirmatory analyses

The results are summarized in Figure A4.1. Threat was negatively associated with performance on the Rotation Span Task ($\beta = -0.13$, $p = .014$), Operation Span Task ($\beta = -0.14$, $p = .014$), and binding trials of the Binding-Updating Task ($\beta = -0.12$, $p = .014$). Unpredictability in perceived scarcity was positively associated with performance on the Rotation Span Task ($\beta = 0.13$, $p = .014$). None of the types of adversity were significantly associated with performance on the updating trials of the Binding-Updating Task.

In addition, there was some limited evidence for practical equivalence, especially for unpredictability in the income-to-needs ratio, which showed a practically equiva-

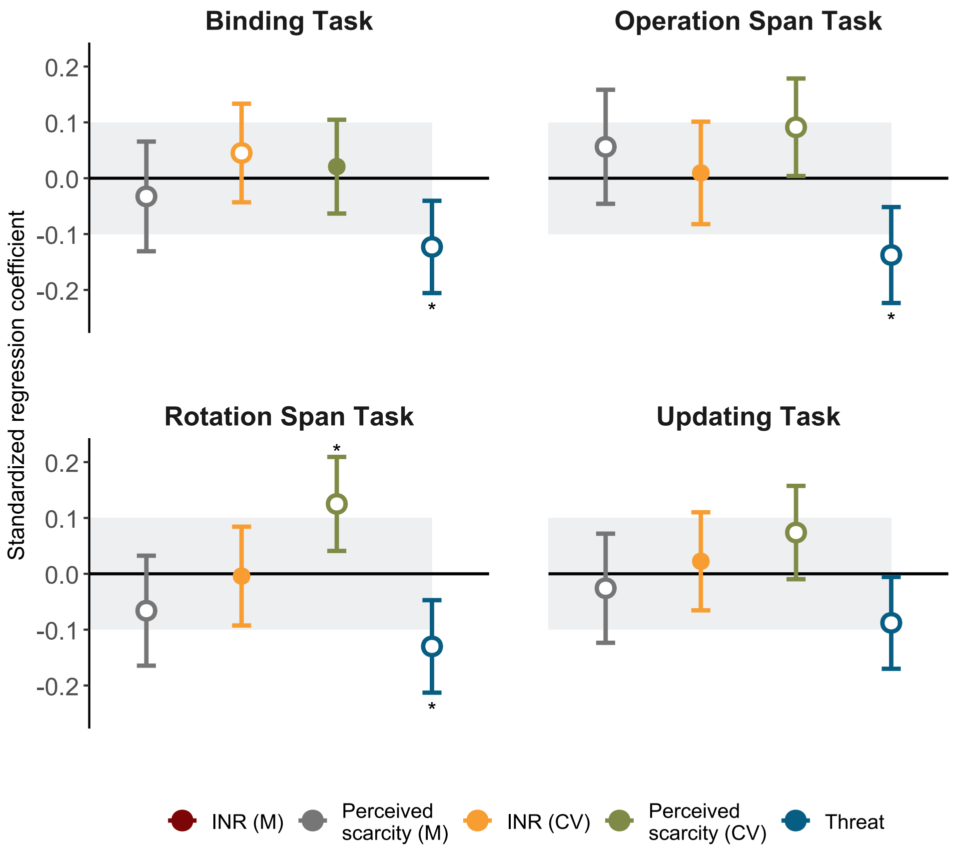



**Figure A4.1.** Exploratory (non-preregistered) results testing the association between threat, deprivation, and unpredictability on residual variances of separate WM tasks. The gray area shows the area of practical equivalence. Solid points indicate effects outside the area of practical equivalence, which was true for all effects. Standard errors represent the 95% confidence intervals. CV = coefficient of variation; INR = income-to-needs ratio; M = mean; WM = working memory.

lent association with the Operation Span Task, Rotation Span Task, and Updating Task. We also found a practically equivalent association between unpredictability in perceived scarcity and the Binding Task.

### Bayes Factors for equivalence tests

As a robustness check, we calculated Bayes factors for the preregistered equivalence tests using the *bain* package (Hoijtink et al., 2019), in which we evaluated evidence in favor of the hypothesis that the effects fell within the equivalence bounds, relative to the hypothesis that the effects fell outside the equivalence bounds. The results are summarized in Table A4.3. For all but one association, the model comparisons showed at least strong evidence in favor of the hypothesis that the effects fell within the equivalence bounds ($BF_{10}$ ranging between 16.9 and 158.9. The only exception was the association between threat and WM capacity, for which we found moderate evidence for the hypothesis that the effect fell within the equivalence bounds ($BF_{10}$ = 5.5. Thus, these results were inconsistent with the preregistered frequentist equivalent tests.

**Table A4.3.** Bayes Factors for practical equivalence tests.

| Hypothesis | $BF_{10}$ |
|---|---|
| -0.1 < (WM capacity ~ INR CV) < 0.1 | 31.6 |
| -0.1 < (WM capacity ~ INR mean) < 0.1 | 84.3 |
| -0.1 < (WM capacity ~ Perc. scarcity CV) < 0.1 | 158.9 |
| -0.1 < (WM capacity ~ Perc. scarcity mean) < 0.1 | 37.7 |
| -0.1 < (WM capacity ~ Threat) < 0.1 | 5.5 |
| -0.1 < (WM updating ~ INR CV) < 0.1 | 91.7 |
| -0.1 < (WM updating ~ INR mean) < 0.1 | 16.9 |
| -0.1 < (WM updating ~ Perc. scarcity CV) < 0.1 | 158.9 |
| -0.1 < (WM updating ~ Perc. scarcity mean) < 0.1 | 40.2 |
| -0.1 < (WM updating ~ Threat) < 0.1 | 77.0 |

BF = Bayes factor; CV = coefficient of variance, INR = income-to-needs ratio, M = mean, Perc. Scarcity = perceived scarcity, WM = working memory

## Section 3. Histograms of independent measures

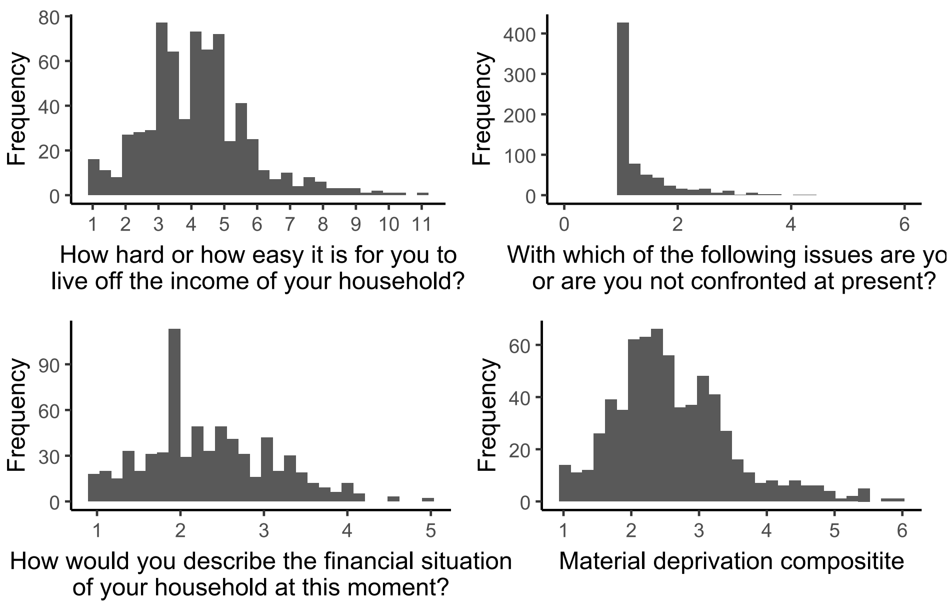



**Figure A4.2.** Histograms for mean perceived material deprivation over time.

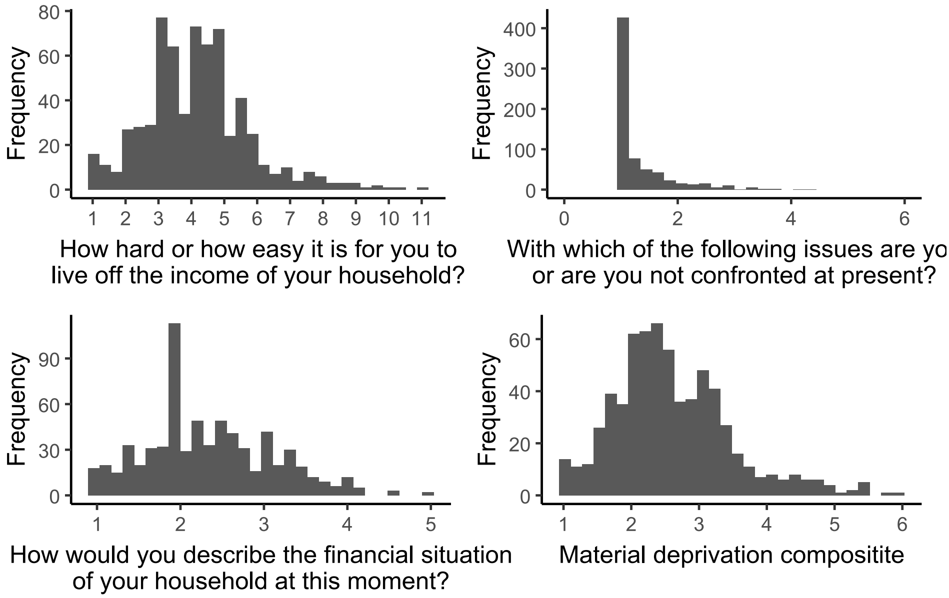



**Figure A4.3.** Histogram for the mean income-to-needs ratio over time.

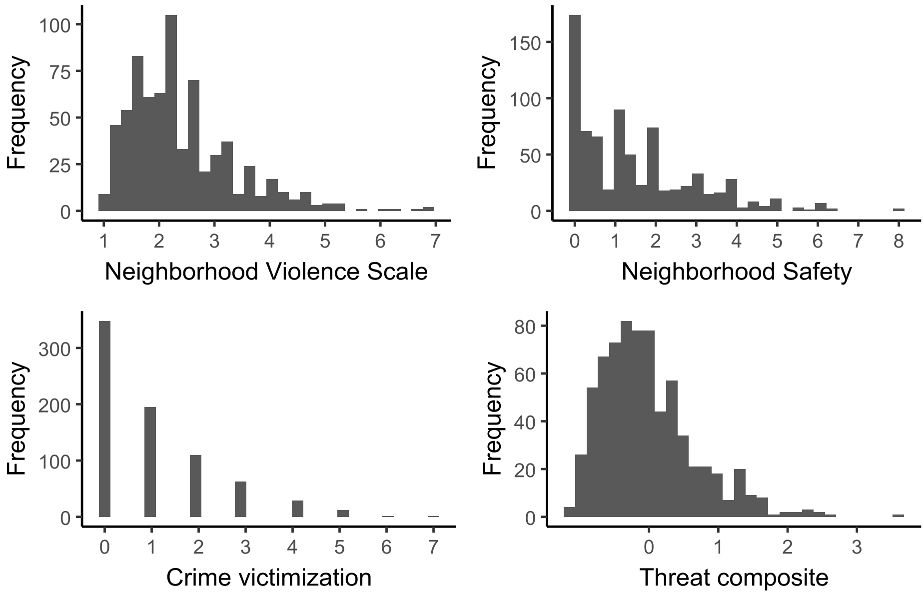



**Figure A4.4.** Histograms for mean threat exposure over time.

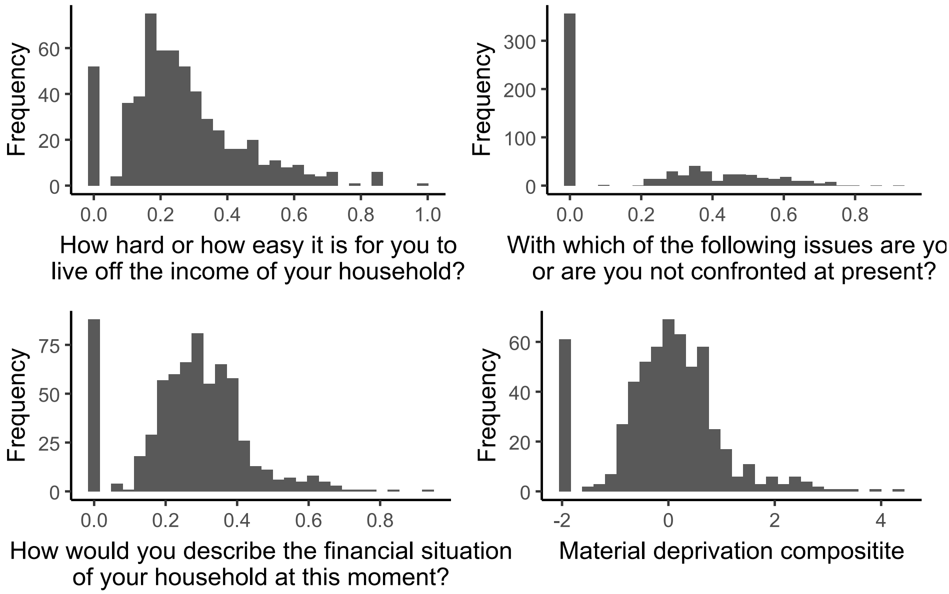



**Figure A4.5.** Histograms for unpredictability in measures of perceived scarcity over time (coefficient of variation).

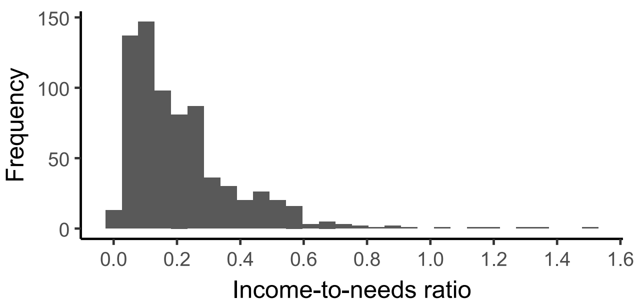



**Figure A4.6.** Histograms for unpredictability in the income-to-needs ratio over time (coefficient of variation).

## Nederlandse samenvatting

Decennia aan onderzoek tonen aan dat mensen uit chronisch stressvolle omgevingen gemiddeld genomen slechter presteren op cognitieve taken. Dit is in verband gebracht met beperkingen in verschillende cognitieve vaardigheden. Deze bevindingen hebben geleid tot een proliferatie van *deficit modellen*, die de nadruk leggen op de negatieve gevolgen van stressvolle omstandigheden op de sociale en cognitieve ontwikkeling. In recenter onderzoek is echter steeds meer aandacht voor adaptaties in specifieke vaardigheden en strategieën. Zulke adaptaties kunnen mogelijk leiden tot verbeteringen in vaardigheden die nuttig zijn voor het oplossen van unieke uitdagingen in stressvolle omstandigheden (bijvoorbeeld het herkennen van gevaar of het detecteren van infrequente beloningen). Een betere integratie van deficit- en adaptatieperspectieven is belangrijk voor de ontwikkeling van gebalanceerde wetenschappelijke en maatschappelijke inzichten.

Desondanks is het moeilijk gebleken om deficit- en adaptatieperspectieven met elkaar te integreren. In dit proefschrift beargumenteer ik dat één van de redenen hiervoor methodologisch van aard is. In beide typen onderzoek worden cognitieve vaardigheden vaak gemeten met behulp van ruwe prestaties op cognitieve taken. Voorbeelden van ruwe prestatiematen zijn gemiddelde reactietijden of het percentage correcte responsen. Zo worden langzamere responsen op inhibitietaken doorgaans geïnterpreteerd als een probleem in het negeren van afleidingen. Recent onderzoek in cognitieve psychologie laat echter zien dat het gebruik van ruwe prestatiematen meerdere limitaties behelst. Door deze limitaties overschatten onderzoekers mogelijk de mate waarin specifieke EF vaardigheden worden verslechterd door blootstelling aan stressvolle omstandigheden, en missen ze mogelijk adaptaties in vaardigheden en strategieën.

Dit proefschrift focust op twee limitaties van het gebruik van ruwe prestatiematen in de context van het meten van *executieve functies* (EF). EFs zijn een verzameling cognitieve vaardigheden die betrokken zijn bij plannen, redeneren, en doelgericht gedrag. De eerste limitatie is dat prestaties op EF taken niet alleen worden beïnvloed door specifieke EFs, maar ook door andere cognitieve processen. Voorbeelden van zulke cognitieve processen zijn de mate van voorzichtigheid waarmee iemand beslissingen neemt en de snelheid waarmee iemand een respons uitvoert. Dit betekent dat de ruwe prestaties van twee mensen kunnen verschillen zelfs als ze niet verschillen in hun EF vaardigheid, bijvoorbeeld wanneer de één relatief voorzichtiger is dan de ander. De tweede limitatie is dat prestaties op EF taken niet alleen worden beïnvloed door specifieke EF vaardigheden, maar ook door algemene processen die van invloed zijn op verschillende EF taken. Als een dergelijk algemeen cognitief proces leidt tot lagere prestaties op meerdere taken kan dit verkeerd worden geïnterpreteerd als meerdere verslechterde vaardigheden. Beide limitaties staan een integratie van deficit- en adaptatiemodellen in de weg (**Hoofdstuk 1**).

Om de eerste limitatie te addresseren maak ik in dit proefschrift gebruik van cognitief modelleren, specifiek het *Drift Diffusion* Model (DDM). Het DDM geeft een verklaring van de manier waarop mensen beslissingen nemen op taken met binaire keuzeopties, zoals veel EF taken. Het verklaart verschillen in reactiesnelheden en accuratesse als een combinatie van (1) de snelheid waarmee iemand informatie verzamelt (de *drift rate*), (2) de hoeveelheid informatie die iemand nodig heeft om een beslissing te nemen (met andere woorden, hoe voorzichtig iemand is; de *boundary separation*) en (3) de snelheid van processen die niet direct betrokken zijn bij nemen van beslissingen, zoals het encoderen van informatie aan het begin van de taak en het uitvoeren van een actie zodra een beslissing is genomen (de *non-decision time*. Door het DDM toe te passen op empirische data van reactiesnelheden en accuratesse is het mogelijk om individuele schattingen te krijgen van deze drie cognitieve processen. Om de tweede limitatie te addresseren gebruik ik *structural equation modeling*. Deze techniek maakt het mogelijk om te onderzoeken in hoeverre prestaties op EF taken worden beïnvloed door cognitieve processen die worden gedeeld door taken, versus cognitieve processen die uniek zijn voor specifieke taken.

Het onderzoek beschreven in dit proefschrift laat zien dat de samenhang tussen stressvolle omstandigheden en specifieke EF vaardigheden wordt overschat door het gebruik van ruwe prestatiematen. Dit blijkt in de eerste plaats uit het feit dat negatieve associaties tussen blootstelling aan stressvolle omstandigheden en *drift rates* worden gedreven door algemene processen die van invloed zijn op verschillende EF taken. Amerikaanse kinderen met meer blootstelling aan dreiging (maar niet deprivatie) hedden een lagere algemene *drift rate* op drie EF taken en een basale verwerkingssnelheidstaak (**Hoofdstuk 2**). Nederlandse volwassenen met meer blootstelling aan dreiging in de volwassenheid, en zowel dreiging als deprivatie in de kindertijd, hebben eveneens een lagere algemene *drift rate* (**Hoofdstuk 3**). Tot slot presteren jongvolwassenen met meer blootstelling aan dreiging en onvoorspelbaarheid in de kindertijd slechter op de Flankertaak vanwege lagere perceptuele verwerking, niet vanwege een lagere inhibitievaardigheid (**Hoofdstuk 4**). In **Hoofdstuk 6** bespreek ik verschillende mogelijke interpretaties van dit algemene proces. De meest waarschijnlijke verklaring—op basis van de huidige bevindingen en voorgaand onderzoek—is dat het een indicatie is van een lagere algemene verwerkingssnelheid. Het is echter mogelijk dat de algemene *drift rate* factor een combinatie is van verschillende processen, zoals iemands huidige staat (motivatie, vermoeidheid, etc.) en meer stabiele individuele eigenschappen.

Na een correctie voor algemene processen is er weinig tot geen bewijs dat blootstelling aan stressvolle omstandigheden samenhangt met specifieke EF vaardigheden. Amerikaanse kinderen met meer blootstelling aan dreiging hebben taakspecifieke *drift rates* die praktisch equivalent zijn aan die van kinderen uit een veiligere omgeving (**Hoofdstuk 2**). Dit suggereert dat hun EF vaardigheden even goed ontwikkeld zijn. Deze studie bevatte echter maar één taak per EF vaardigheid; hierdoor was het niet mogelijk om te onderzoeken in hoeverre taakspecifieke *drift rates* een reflectie waren van specifieke EF vaardigheden. In een vervolgstudie met Nederlandse volwassenen includeer ik meerdere taken per EF vaardigheid (inhibitie en aandachtsverschuiving), wat het mogelijk maakt om de EF vaardigheden op latent niveau te schatten. Na een correctie voor algemene processen is de overgebleven variantie echter niet toe te schrijven aan specifieke EF vaardigheden. Hoewel sommige vormen van stressvolle omstandigheden (met name dreiging in de kindertijd) negatief samenhangen met *drift rates*, zijn al deze effecten onafhankelijk van elkaar, ook voor taken die dezelfde vaardigheid beogen te meten. Met andere woorden, mensen met meer blootstelling aan stressvolle omstandigheden verwerken informatie op specifieke taken langzamer vanwege unieke aspecten van de taak, niet vanwege verslechterde EF vaardigheden die worden gedeeld door sommige taken. Tot slot hangt blootstelling aan stressvolle omstandigheden ook niet samen met de vaardigheid om informatie in het werkgeheugen te updaten, nadat ik corrigeer voor de capaciteit van het werkgeheugen (**Hoofdstuk 5**).

Tot slot laat mijn onderzoek zien dat mensen uit meer stressvolle omstandigheden in sommige gevallen andere strategieën gebruiken dan mensen uit minder stressvolle omstandigheden. Amerikaanse kinderen met meer blootstelling aan dreiging hebben een (al dan niet bewuste) strategie om langer de tijd te nemen en daarmee hun accuratesse te vergroten (**Hoofdstuk 2**). Ik vind echter geen bewijs voor een dergelijke strategie bij jongvolwassenen (**Hoofdstuk 4**) en volwassenen (**Hoofdstuk 3**). Daarnaast suggereert mijn onderzoek ook dat jongvolwassenen met meer blootstelling aan dreiging en onvoorspelbaarheid in de kindertijd langzamer zijn in het verwerken van informatie op de Flankertaak omdat ze een meer holistische informatieverwerkingsstijl hebben (**Hoofdstuk 4**). Dit betekent dat ze meer geneigd zijn om te focussen op globale eigenschappen van visuele informatie in plaats van individuele stimuli. Mogelijk is dit een adaptatie die mensen beter in staat stelt om perifere informatie snel te detecteren.

Dit proefschrift biedt een aantal aanbevelingen voor toekomstig wetenschappelijk onderzoek (**Hoofdstuk 6**). Het is belangrijk om rekening te houden met het feit dat (1) prestaties op EF taken worden beïnvloed door meerdere cognitieve processen en (2) dat prestaties op EF taken voor een substantieel deel worden beïnvloed door algemene processen die van invloed zijn op meerdere taken, niet door specifieke EF vaardigheden. Daarnaast zouden studies idealiter twee of meer taken per EF vaardigheid includeren, zodat deze vaardigheden op latent niveau geschat kunnen worden. De combinatie van cognitief modelleren en *structural equation modeling* kan een centrale rol spelen in twee huidige onwikkelingen in het veld. Ten eerste stelt het toekomstig onderzoek beter in staat om specifieke hypothesen te testen over de interactie van deficit- en adaptatieprocessen binnen dezelfde persoon. Ten tweede kan het helpen in het beter begrijpen waarom mensen uit stressvolle omstandigheden gevoeliger lijken te zijn voor verschillende typen inhoud van cognitieve taken, en welke cognitieve processen hiervoor verantwoordelijk zijn. Op deze manier is cognitief modelleren een onmisbare techniek voor de volgende generatie onderzoekers in dit veld.

# Acknowledgements

Some major life decisions are based on meticulous planning. Others just sort of come to you out of nowhere while driving home from work. I tend to like the latter ones better. The decision that ultimately led to starting this PhD definitely falls in this category. In that sense, the first acknowledgement should really go out to Supertramp. For some reason that I do not claim to understand, listening to Crime of the Century on that drive back from Eindhoven on 13 December 2019 sparked some weird neural activation that led me to start pursuing a PhD again.

Willem, we first met all the way back in January 2017 when I was still in the first year of my Master. That day, in true Willem fashion, we walked all the way from Utrecht Central Station to the Uithof, talking about science the whole way. After that first meeting, it took nearly three years for us to reconnect and embark on this PhD. Looking back, I can honestly say that I truly enjoyed (almost) every day of the past four years, for a large part thanks to your contagious enthusiasm, stupid humor, and genuine kindness. I am really grateful for all the cool opportunities you provided, and I look forward to continuing our collaboration in the future. I am sorry you were never able to get me to write “complementary” correctly (you will get a second change during my postdoc, but don’t get your hopes up). Jean-Louis, I really liked having you on my team. Even though the topic of my dissertation ended up venturing somewhat outside of your area of expertise, I could always count on thoughtful feedback and discussions. I thoroughly enjoyed our visits to Freiburg; I felt immediately at home at the MPI, and I really enjoyed organizing the Open Science workshop. Jaap, your time as official promotor was relatively short, but of course you were already involved as a co-author before then. Thank you for your for your valuable input, and for regularly checking in on me after Willem and Ethan left Utrecht University.

Ethan, I feel very fortunate to have had you on board as a co-promotor, and effectively as daily supervisor throughout most of my PhD. I really value your critical but pragmatic, thorough but no-nonsense way of working. It was especially nice to nerd out with you over new R packages, ways of further optimizing our project workflows, the multiverse, as well as complaining about the *baggerzooi* that academia can be at times. This dissertation would have looked very differently without your help in finding my own voice.

Anna-Lena, reaching out to you early on was one of the best decisions that I made during my PhD. It led to your extensive involvement in three of my four projects, altering their trajectory for the better. My visits to Mainz, in particular the month-long visit in fall 2023, were absolute highlights of my PhD. You were a promotor in all but name (due to strict UU regulations), which I am very grateful for. I am also thankful to everyone in your lab for being so welcoming. Jan, Simon, Christoph, Wanja, Meike, & Jean-Paul, thank you for all the nice discussions and after-work beers. A special thank you to Meike and Jean-Paul for so warmly welcoming me into your home during several boardgame nights. I feel really fortunate for having met all of you.

Meriah, It was great having you involved in two of my dissertation chapters. Your extensive knowledge on the topic coupled with our shared interest in cognitive modeling made you a great addition to the team. A huge thanks as well for hosting me at Stanford! I hope we will continue to collaborate on cool projects in the future.

I would also like to thank all my (ex) OWP colleagues for creating such a warm, supportive, and inspiring work environments. Despite starting my PhD in the middle of the second Covid wave, and therefore not meeting most of you in person until quite some time later, I felt right at home. I massively enjoyed all the lunch chats, writing weeks, drinks, Christmas dinners, and outings. A special thanks to my fellow (ex) Jonkies, Anouk, Berfu, Danni, Eva, Gaby, Jenna, Judith, Lisa, Mengtian, Nora, Saira, Sophie, Sun, Thao, Xiaohuan, Yixin, Daniëlle, Isa, Isabelle, Sasha, and Tessa. Jenna, Sun, I really enjoyed sharing H2.56 with you. Even on days when the rest of the hallway was mostly deserted, H2.56 was guaranteed to be filled with people. Sun, all my office plants would have died a thousand times over if not for your loving care during my holidays and conferences. Jenna, Judith, thank you for agreeing to be my paranymphs (Let's hope the baby will have the patience to wait until after April 16th!).

Aside from my OWP colleagues, I was also fortunate to meet and interact with my MPI colleagues on several occasions. Sebas, Benoît, Aniek, Tim, Shaina (and many others), thank you for sharing your office with me and showing us around in your beautiful city. You all made me feel at home from the very first visit. Freiburg will always have a special place in my heart thanks to you.

Jan, thank you for sitting next to me in Tenerife while I was writing these acknowledgements.

I also want to thank my family, especially my parents and Marit. If there is one thing I have learned from working on this topic, it is just how vital a warm and supportive family environment is—and how easy it is to take it for granted. I was fortunate to grow up surrounded by that warmth and support, and I still feel it today. My work probably stopped making a lot of sense to you a long time ago. Regardless, it is largely because of your unconditional support over the last 32 years that I got this far.

Tolstoy, Nebula, & Novi, your support over the past four years—very much conditional on food and cuddles—cannot be overstated. Tolstoy, thank you for always barging in just when I'm about to leave for work. I would have been way too punctual if not for you. Nebula, thank you for consistently getting up with me in the mornings, and for the occasional nightly present (you really shouldn't… no, really). Novi, you're the embodiment of tough love. You don't ask for cuddles, you enforce them. You have found a willing victim in me.

Anaïs, while Supertramp may have provided the first spark, you were the one to really give me the necessary push into a new adventure. Finishing this PhD would have been a lot more difficult and a lot less fun without you. Career-wise, it has been a solid decision to marry the smartest person I know. You are there to cheer me on when things are going well, and you are there to support me when I'm down. On top of it all, you spend hours creating awesome drawings for my book. As the Editor of the Journal of Husband Science, I could never thank you enough for everything that you do for me.

## About the author

Stefan was born in Gouda, The Netherlands, on 27 February 1993. In 2016, he obtained his Bachelor's degree (with Honours) at Utrecht University, with a specialization in social psychology. During his Bachelor's, Stefan developed a fascination with cognitive processes and their interactions with the environment. He completed a research internship at Utrecht University in 2015 in which he investigated how cognitive simulations of food consumption induce salivation. He became co-author on the resulting article, and continued the project in his Bachelor thesis, manipulating cognitive simulations using mindfulness.

In 2018, Stefan obtained his Master's degree in Social and Health Psychology (research master, cum laude) at Utrecht University. During his Master's, he completed an internship using agent-based modeling to investigate how inferences of trustworthiness based on facial features develop over evolutionary time. His Master thesis investigated whether monetary cues can unconsciously affect cognitive effort.

After obtaining his Master's degree, Stefan started working as a junior researcher at Philips, Eindhoven (The Netherlands) in 2018. He worked on the development of a novel digital neuropsychological assessment platform to assess cognitive functioning in older adults. Specifically, he was responsible for translating test performance into cognitive domain scores, computing normed scores, and test validation. His time at Philips further consolidated his interest in cognitive processes. Aside from further consolidating his fascination for cognitive processes, Stefan also discovered a strong interest in programming. In 2021, Stefan decided to leave Philips in order to pursue a PhD at Utrecht University and the Max Planck Institute for the Study of Crime, Security, and Law. His PhD project focused on how exposure to adversity shapes executive functioning abilities, using cognitive modeling to better understand how cognitive deficits and adaptations may operate within the same individual. He went on research visits to Mainz University (Germany) and Stanford University (USA). Alongside his research, Stefan also developed and taught workshops on Multiverse analysis, open science, and creating reproducible projects in R. Stefan currently works as a postdoc at the University of Amsterdam.

In his free time, Stefan enjoys going on overly long cycling trips and spending time with his friends, family, wife, and three cats.

# Publications

## Peer-reviewed articles

**Vermeent, S.**, Schubert, A.-L., DeJoseph, M.L., Denissen, J.J.A., van Gelder, J.-L., & Frankenhuis, W.E. (2025). Inconclusive evidence for associations between adverse experiences in adulthood and working memory performance. *Royal Society Open Science, 12*(1), 241837.

**Vermeent, S.**, Young, E.S., DeJoseph, M.L., Schubert, A.-L., & Frankenhuis, W.E. (2024). Cognitive deficits and enhancements in youth from adverse conditions: An integrative assessment using Drift Diffusion Modeling in the ABCD study. *Developmental Science, 27*(4), e13478. https://doi.org/10.1111/desc.13478

**Vermeent, S.**, Young, E.S., van Gelder, J.-L., & Frankenhuis, W.E. (2024). Childhood adversity is not associated with lowered inhibition, but slower perceptual processing: A Drift Diffusion Model analysis. *Cognitive Development, 71*, 101479. https://doi.org/10.1016/j.cogdev.2024.101479

Young, E.S., **Vermeent, S.**, Frankenhuis, W.E., Nivison, M.D., Simpson, J.A., & Roisman, G.I. (2024). How does adversity relate to performance across different abilities within individuals. *Development and Psychopathology*, 1-18. https://doi.org/10.1017/S0954579424001433

Klaming, L., Spaltman, M., **Vermeent, S.**, van Elswijk, G., Miller, J.B., & Schmand, B. (2024). Test-retest reliability and reliable change index of the Philips IntelliSpace Cognition digital test battery. *The Clinical Neuropsychologist, 38*(7), 1707-1725. https://doi.org/10.1080/13854046.2024.2315747

**Vermeent, S.**, Spaltman, M., van Elswijk, G., Miller, J.B., & Schmand, B. (2022). Philips IntelliSpace Cognition digital test battery: Equivalence and measurement invariance compared to traditional analog test versions. *The Clinical Neuropsychologist, 36*(8), 2278-2299. https://doi.org/10.1080/13854046.2021.1974565

**Vermeent, S.**, Dotsch, R., Schmand, B., Klaming, L., Miller, J.B., & van Elswijk, G. (2020). Evidence of validity for a newly developed digital cognitive test battery. *Frontiers in Psychology, 11*, 770. https://doi.org/10.3389/fpsyg.2020.00770

Keesman, M., Aarts, H., **Vermeent, S.**, Häfner, M., & Papies, E.K. (2016). Consumption simulations induce salivation to food cues. *PLoS ONE, 11*(11), e0165449. https://doi.org/10.1371/journal.pone.0165449

### Manuscripts under review

**Vermeent, S.**, Schubert, A.-L., & Frankenhuis, W.E. (under review). Adversity is associated with lower general processing speed rather than specific executive functioning abilities. https://doi.org/10.31234/osf.io/kqhf7

### Book chapters

Custers, R., **Vermeent, S.**, & Aarts, H. (2019). Does goal pursuit require conscious awareness? In R.M. Ryan (Ed.), Oxford Handbook of Motivation. New York: Oxford University Press, Inc.