# Development of Inhibitory Control in Kindergarten Children

Inaugural-Dissertation

zur Erlangung der Doktorwürde der Fakultät für Humanwissenschaften der

Julius-Maximilians-Universität Würzburg

Vorgelegt von

Daniel Schulz

aus Nürnberg

Würzburg 2023

| | |
|---|---|
| Erstgutachter: | PD Dr. Eva Michel |
| Zweitgutachter: | PD Dr. Simon Tiffin-Richards |
| Tag des Kolloquiums: | 20.02.2024 |

# Table of Contents

**Summary ........................................................................................ 4**

**Chapter I: Theoretical and Empirical Background ........................... 8**

**Chapter II: Present Research ..................................................... 37**

**Chapter III: Study 1 .............................................................. 56**

Schulz, D., Richter, T., Schindler, J., Lenhard, W. & Mangold, M. (2023). Using accuracy and response times to assess inhibitory control in kindergarten children: An analysis with explanatory item response models. *Journal of Cognition and Development, 24*(1), 82-104. https://doi.org/10.1080/15248372.2022.2119977

**Chapter IV: Study 2 ............................................................. 99**

Schulz, D., Segerer, R., Lenhard, W., Mangold, M., Schindler, J. & Richter, T. (2023). Assessing inhibitory control in kindergarten children: Validity of integrating response accuracy and response latency. *Cognitive Development, 68*, Article 101392. https://doi.org/10.1016/j.cogdev.2023.101392

**Chapter V: Study 3 .............................................................. 145**

Schulz, D., Lenhard, W., Mangold, M., Schindler, J. & Richter, T. (submitted). Balancing accuracy and speed in the development of inhibitory control. *Manuscript submitted for publication*,

**Chapter VI: General Discussion ............................................ 177**

**Acknowledgements .............................................................. 206**

**Summary**

This dissertation explores the development and assessment of inhibitory control – a crucial component of executive functions – in young children. Inhibitory control, defined as the ability to suppress inappropriate responses (Verbruggen & Logan, 2008), is essential for adaptable and goal-oriented behavior. The rapid and non-linear development of this cognitive function in early childhood presents unique challenges for accurate assessment**.** As children age, they often exhibit a ceiling effect in terms of response accuracy (Petersen et al., 2016), underscoring the need to consider response latency as well. Ideally, combining response latency with accuracy could yield a more precise measure of inhibitory control (e.g., Magnus et al., 2019), facilitating a detailed tracking of developmental changes in inhibitory control across a wider age spectrum. The three studies of this dissertation collectively aim to clarify the relationship between response accuracy, response latency, and inhibitory control across different stages of child development. Each study utilizes a computerized Pointing Stroop Task (Berger et al., 2000) to measure inhibitory control, examining the task's validity and the integration of dual metrics for a more comprehensive evaluation.

The first study focuses on establishing the validity of using both response accuracy and latency as indicators of inhibitory control. Utilizing the framework of explanatory item-response modeling (De Boeck & Wilson, 2004), the study revealed how the task characteristics congruency and item position influence both the difficulty level and timing aspects in young children’s responses in the computerized Pointing Stroop task. Further, this study found that integrating response accuracy with latency, even in a basic manner, provides additional insights. Building upon these findings, the second study investigates the nuances of integrating response accuracy and latency, examining whether this approach can account for age-related differences in inhibitory control. It also explores whether response latencies may contain different information depending on the age and proficiency of the children. The study

leverages novel and established methodological perspectives to integrate response accuracy and latency into a single metric, showing the potential applicability of different approaches for assessing inhibitory control development. The third study extends the investigation to a longitudinal perspective, exploring the dynamic relationship between response accuracy, latency, and inhibitory control over time. It assesses whether children who achieve high accuracy at an earlier age show faster improvement in response latency, suggesting a non-linear maturation pathway of inhibitory control. The study also examines if the predictive value of early response latency for later fluid intelligence is dependent on the response accuracy level.

Together, these empirical studies contribute to a more robust understanding of the complex interaction between inhibitory control, response accuracy, and response latency, facilitating valid evaluations of cognitive capabilities in children. Moreover, the findings may have practical implications for designing educational strategies and clinical interventions that address the developmental trajectory of inhibitory control. The nuanced approach advocated in this dissertation suggests prioritizing accuracy in assessment and interventions during the early stages of children's cognitive development, gradually shifting the focus to response latency as children mature and secure their inhibitory control abilities.

**Zusammenfassung**

Die vorliegende Dissertation erforscht die Erfassung und Entwicklung von Inhibitionskontrolle bei jungen Kindern – einer zentralen Komponente der Exekutiven Funktionen. Inhibitionskontrolle, also die Fähigkeit, automatisierte aber unangemessene Reaktionen zu unterdrücken (Verbruggen & Logan, 2008), ist wesentlich für adaptives und zielgerichtetes Verhalten. Die schnelle und nichtlineare Entwicklung dieser kognitiven Funktion im frühen Kindesalter gestaltet eine präzise Messung herausfordernd. Mit zunehmendem Alter der Kinder zeigt sich häufig ein Deckeneffekt hinsichtlich der Antwortgenauigkeit (Petersen et al., 2016), was die Notwendigkeit hervorhebt, auch die Reaktionszeit in Betracht zu ziehen. Idealerweise könnte durch die Integration von Reaktionszeit und Antwortgenauigkeit ein Messwert berechnet werden (z.B. Magnus et al., 2019), welcher eine detaillierte Erfassung von Entwicklungsveränderungen der Inhibitionskontrolle über ein breiteres Altersspektrum hinweg ermöglicht. Die drei Studien dieser Dissertation zielen darauf ab, die Beziehung zwischen Antwortgenauigkeit, Reaktionszeit und Inhibitionskontrolle in verschiedenen Stadien der kindlichen Entwicklung zu untersuchen. Jede Studie nutzt eine computergestützte Inhibitionsaufgabe, den computerized Pointing-Stroop Task (cPST; Berger et al., 2000), um die Inhibitionskontrolle zu messen, wobei die Validität dieses Tests und die Integration von Antwortgenauigkeit und Reaktionszeit für eine umfassendere Bewertung untersucht werden.

In der ersten Studie wird untersucht, ob sowohl Antwortgenauigkeit als auch Reaktionszeit valide Indikatoren für Inhibitionskontrolle in jungen Kindern darstellen. Unter Verwendung von explanatorischen Item-Response-Modellen zeigte die Studie, wie die Aufgabenmerkmale Kongruenz und Item-Position die Aufgabenschwierigkeit sowohl in Bezug auf Antwortgenauigkeit als auch Reaktionszeit im cPST beeinflussen. Darüber hinaus zeigten sich erste Hinweise, dass bereits eine rudimentäre Integration von

Antwortgenauigkeit und Reaktionszeit zusätzliche Einsichten liefert. Aufbauend auf diesen Erkenntnissen untersucht die zweite Studie die Feinheiten der Integration von Antwortgenauigkeit und Reaktionszeit und prüft, ob moderne Methoden der Integration dieser beiden Metriken altersbedingte Unterschiede in der Inhibitionskontrolle berücksichtigen können. Sie erforscht auch, ob sich aus den Reaktionszeiten in Inhibitionsaufgaben, abhängig vom Alter und Können der Kinder, unterschiedliche Schlussfolgerungen ziehen lassen. Die Studie nutzt neue und etablierte methodische Ansätze, um Antwortgenauigkeit und Reaktionszeit zu einer Metrik zu integrieren und zeigt die potenzielle Anwendbarkeit verschiedener Ansätze zur Bewertung der Entwicklung der Inhibitionskontrolle. Die dritte Studie erweitert die Untersuchung auf eine Längsschnittperspektive und erforscht die dynamische Beziehung zwischen Antwortgenauigkeit, Reaktionszeit und Inhibitionskontrolle im Laufe der Entwicklung. Sie betrachtet, ob Kinder, die in jüngerem Alter eine hohe Genauigkeit erreichen, eine schnellere Verbesserung in der Reaktionszeit zeigen. Die Studie untersucht weiter, ob der prädiktive Wert von Reaktionszeit für zukünftige fluide Intelligenz in Abhängigkeit zu der Antwortgenauigkeit steht.

Zusammen tragen diese empirischen Arbeiten zu einem tieferen Verständnis der komplexen Interaktion zwischen Inhibitionskontrolle, Antwortgenauigkeit und Reaktionszeit bei und erleichtern valide Bewertungen dieser kognitiven Fähigkeiten bei Kindern. Darüber hinaus könnten die Ergebnisse praktische Implikationen für die Gestaltung von Interventionen haben, die sich mit dem Entwicklungsverlauf der Inhibitionskontrolle befassen. Der in dieser Dissertation vertretene Ansatz legt nahe, Antwortgenauigkeit bei der Bewertung und Interventionen während der frühen Phasen der kognitiven Entwicklung von Kindern zu priorisieren und den Fokus allmählich auf die Reaktionszeit zu verlagern, sobald Kinder ihre Inhibitionskontrolle festigen und ausbauen.

# Chapter I

# Theoretical and Empirical Background

## Executive Functions

To fully understand the significance and intricacies of inhibitory control, it is essential first to examine its broader context within the domain of executive functions. Executive functions are effortful and essential cognitive processes responsible for regulating and directing mental activities and behaviors. They enable individuals to pursue goal-directed behavior, adapt to evolving environments, and perform intricate tasks, making them crucial to human life. Despite their fundamental importance, there is some disagreement about the exact nature of executive functions (see Müller & Kerns, 2015).

Diamond (2013) posits three fundamental executive functions: inhibitory control, working memory, and cognitive flexibility. Inhibitory control pertains to the capacity to regulate one's attention, behavior, thoughts, and emotions to override internal inclinations or external temptations (Diamond, 2013). Working memory involves retaining information in the mind and mentally manipulating it (Baddeley & Hitch, 1994; Smith & Jonides, 1999). Cognitive flexibility is the ability to alter perspectives spatially or interpersonally and to adapt to shifting demands or priorities (Davidson et al., 2006; Garon et al., 2008).

Similarly, Miyake et al. (2000) proposed a conceptualization of executive functions, focusing on three primary functions: shifting between tasks or mental sets (Shifting), inhibition of prepotent responses (Inhibition), and updating and monitoring of working memory representations (Updating). Shifting relates to the ability to alternate between multiple tasks or mental sets (Monsell, 1996). Inhibition, akin to Diamond's definition, involves the capacity to intentionally suppress dominant, automatic, or prepotent responses when necessary (Logan & Cowan, 1984). Updating requires monitoring and encoding incoming information for task relevance and subsequently revising items held in working memory by replacing outdated information with new, pertinent information (Jonides & Smith, 1997; Morris & Jones, 1990).

While both Diamond (2013) and Miyake et al. (2000) agree on the importance of inhibition and working memory as critical components of executive functions, they differ in their conceptualizations of the third component. Diamond (2013) highlights cognitive flexibility as the third core executive function, whereas Miyake et al. (2000) emphasize shifting between tasks or mental sets. These distinctions can be viewed as complementary rather than opposing. Cognitive flexibility involves adjusting to changing demands or priorities, while task shifting refers to the ability to switch between multiple tasks, operations, or mental sets. Both perspectives recognize the necessity for individuals to adapt and adjust. Inhibitory control is particularly integral to the workings of executive functions, operating directly or together with other cognitive processes to manage and direct thought and action (Bjorklund & Harnishfeger, 1990). Its value extends beyond mere cognitive regulation, pivotal in various everyday life scenarios (e.g., Mamrot & Hanć, 2019).

I will explore inhibitory control in detail, examining its multifaceted nature, developmental progression, and broad impact on cognitive, social, and emotional domains.

## Inhibitory Control

Inhibitory control fundamentally involves regulating thoughts, emotions, and actions by filtering out irrelevant or unwanted stimuli (e.g., Harnishfeger & Bjorklund, 1993; Simpson & Carroll, 2019). It is vital to adaptive functioning, enabling individuals to concentrate on pertinent information, resist distractions, and maintain goal-oriented behaviors. This capability is critical for a range of cognitive abilities and closely linked to multiple facets of human experience and behavior.

Distinguishing between different types of inhibition is essential for understanding its diverse manifestations. Nigg's (2000) taxonomy outlines four distinct types of inhibition: interference control, cognitive inhibition, behavioral inhibition, and oculomotor inhibition. Interference control involves the suppression of competing stimuli or distractors, as seen in priming and flanker tasks (Gratton et al., 1992). Cognitive inhibition, studied through attentional-orienting paradigms (Posner & Petersen, 1990), focuses on suppressing irrelevant information in working memory. Behavioral inhibition, exemplified by negative priming (Tipper, 1985), concerns inhibiting prepotent responses. Lastly, oculomotor inhibition, investigated through antisaccade and visual-delayed response tasks, involves inhibiting reflexive eye movements. These types underscore the varied nature of inhibitory processes in cognitive functioning.

Harnishfeger (1995) differentiates exclusively between cognitive and behavioral inhibition. Cognitive inhibition involves managing mental contents or processes, which can be intentional and conscious or unintentional and unavailable for conscious introspection (Harnishfeger & Bjorklund, 1993). Instances of cognitive inhibition include suppressing thoughts, clearing incorrect inferences from memory (e.g., Hamm & Hasher, 1992), and managing the removal of irrelevant information from working memory during memory processing (Harnishfeger & Bjorklund, 1993). Behavioral inhibition, in contrast, involves managing overt behavior, such as resisting temptation or inhibiting motor actions (e.g., Mischel et al., 1989). Although Harnishfeger (1995) distinguishes between cognitive and behavioral inhibition, she acknowledges that the two may be related. For example, children may use cognitive inhibition to facilitate behavioral inhibition (e.g., Mischel et al., 1989).

In discussing behavioral inhibition, it is essential to distinguish between 'hot' and 'cold' inhibitory control (e.g., Hao, 2017). Hot inhibitory control involves the affective and motivational aspects of inhibitory capacities, while cold inhibitory control pertains to abstract, non-emotional problem-solving (Zelazo & Müller, 2002). In simpler terms, hot inhibitory control concerns emotional and motivational processes, while cold inhibitory control operates in a cognitive context without emotional engagement. For instance, the Stroop task exemplifies cold inhibitory control, where participants must suppress the prepotent response of reading a word to state the ink's color (Stroop, 1935). Conversely, the delay of gratification task represents hot inhibitory control, requiring the postponement of immediate action for a later reward (e.g., Traverso et al., 2015).

Further distinctions have been proposed, such as inhibitory strength and endurance (Simpson & Carroll, 2019), with strength defined as the ability to resist intense impulses and endurance as the ability to maintain this resistance over time. Even though these distinctions provide a more nuanced understanding of behavioral inhibitory control, the concept remains somewhat undifferentiated in its use in the literature. It is critical to recognize the need for age-fair and consistent assessment techniques to accurately investigate and interpret the role of inhibitory control amidst its various correlates. Researchers must consider the developmental progression of inhibitory control – that is, how the ability to inhibit behaviors or cognitions evolves as children mature – since this can have profound implications on the effectiveness of assessments and the interpretation of findings.

Assessing inhibitory control accurately in research settings involves narrowing down the cognitive or behavioral inhibition one investigates, using age-appropriate methodologies that adapt to children's rapidly evolving cognitive abilities. This approach underscores the importance of using tools sensitive enough to account for developmental stages and cognitive maturation, ensuring assessments yield informative insights into children's inhibitory control capabilities. By employing assessment techniques that are both age-fair and consistent, we can more effectively investigate the role of inhibitory control and its association with other developmental domains, enhancing our understanding of its impact on children's cognitive and behavioral outcomes.

**Correlates of Inhibitory Control**

Notable correlates of inhibitory control are, for example, intelligence, self-regulation, social-emotional development, and academic success. Inhibitory control is linked to intelligence, both via theoretical rationale and through findings from neurological studies, specifically concerning fluid intelligence, which includes reasoning, abstract thinking, and problem-solving capacities (Horn & Cattell, 1967). Inhibitory control is considered an essential component in problem-solving as it enables the suppression of irrelevant behavior and information (Dempster, 1991). This capability enables disengagement from previously unsuccessful solution attempts and preserves working memory resources for processing relevant information and solving problems (Bjorklund & Harshfeger, 1990). Neurologically, inhibitory control is strongly connected to the functionality of the prefrontal cortex, which is considered crucial for higher-order cognitive functions, including fluid intelligence (Dempster, 1991; Moriguchi & Hiraki, 2013).

Empirical studies have shown varied results regarding the correlation between inhibitory control and fluid intelligence. While some evidence supports a clear positive relationship, using the Stroop task (e.g., Yücel et al., 2012), other studies suggest that when updating and inhibitory control are considered simultaneously as predictors of intelligence, the relationship remains significant for updating but not for inhibitory control (Friedman et al., 2006; Duan et al., 2010). Disparities in these findings could be due to the diversity of inhibitory control tasks utilized across studies or the complexity of capturing inhibitory control, as it varies developmentally and methodologically.

Moreover, the significance of inhibitory control extends beyond cognitive abilities to encompass self-regulation and social-emotional competencies, which are particularly critical during early childhood – a phase marked by rapid development and fine-tuning of cognitive and behavioral patterns. For example, Hao (2017) examined inhibitory control correlations with donating behavior in children during early to middle childhood. The study found that, for second graders, donating behavior was predicted by their performance on the fruit Stroop task (Archibald & Kerns, 1999). However, for sixth graders, donating behavior was predicted by their performance on the delay of gratification task (based on Groppe & Elsner, 2014). Similar patterns have been observed with behavioral adjustment, where children with robust inhibitory control demonstrated improved social skills and fewer behavioral issues (Rhoades et al., 2009). The relational strength of these social factors varied significantly based on differing measures of inhibitory control, such as the Peg Tapping (Diamond & Taylor, 1996) and Day/Night (Gerstadt et al., 1994) tasks. The Peg Tapping task assesses children's ability to inhibit a natural tendency to mimic an action. The Day/Night task measures children's ability to inhibit a natural verbal response tendency to a card displayed. These insights underscore the importance of considering developmental stages in the assessment of inhibitory control and highlight how the choice of assessment methods can influence the outcomes of such investigations.

When considering the impact of inhibitory control on academic achievement, a coherent approach to its assessment becomes even more paramount. Ng et al. (2014) found a substantial association between inhibitory control at age four and growth in early math skills over subsequent years. However, in a meta-analysis, Allan et al. (2014) found that the relationship between inhibitory control and academic skills in preschool and kindergarten children depended on the type of inhibitory control assessed. Notably, the authors found that cold inhibitory control tasks, which are decontextualized and lack emotional or motivational significance, were more related to academic skills than hot inhibitory control tasks, which have emotional or motivational significance (see Brock et al., 2009).

In light of these findings, it is clear that inhibitory control—particularly behavioral inhibition – is a significant factor in development, necessitating further research, consistent assessment, and careful consideration of the developmental context. It is crucial to recognize the various types of inhibition and employ age-fair and developmentally sensitive assessment methods to advance our understanding and develop interventions that will positively influence children's academic, cognitive, and social-emotional outcomes. These assessments enable practitioners to pinpoint and support the development of inhibitory control in ways that consider the child's current capabilities while fostering positive future trajectories. Such conscientious application of robust research methodologies and considerations can illuminate inhibitory control's pivotal role, not just as a singular cognitive construct but within the broader context of human development.

## Development and Assessment of Inhibitory Control

### Approaches to Assessing Inhibitory Control in Children

Measuring inhibitory control, particularly response inhibition, in young children is approached through various methods. However, many assessments in practice prioritize response accuracy but often fail to accurately measure response latency, despite its potential significance as an aspect of inhibitory control (e.g., Magnus et al., 2019).

The Stroop task (Stroop, 1935) is a commonly administered test among adults. In this task, participants are shown color words printed in differently colored inks and are instructed to name the ink color and ignore the printed word. In the conflict condition, the color of the ink is incongruent with the word (e.g., the word "red" printed in blue ink), thus requiring participants to suppress the prepotent response of reading the word and instead identify the color of the ink. This test effectively measures inhibitory control as participants must inhibit the automated response to read the word in favor of stating the ink color, demonstrating the capacity for response inhibition.

Children's assessments are adapted to be engaging and to sustain the child's interest. For example, the Day-Night Test (Gerstadt et al., 1994) presents a playful approach. In this test, children are shown either a bright sun on a white card or a moon with stars on a dark card. They are asked to say "night" when shown the sun card and "day" when presented with the moon and stars card. However, the offline implementation of the Day-Night Test does not precisely measure response latency. Another child-friendly approach is the Silly Sound Stroop task (Willoughby et al., 2011). Children are instructed to say *meow* when shown a picture of a dog and *bark* when shown a picture of a cat, thus requiring the suppression of the automatic response associated with the presented image. As with the Day-Night Test, the Silly Sound Stroop task is traditionally implemented in an offline format that fails to measure response latency accurately.

The challenge, however, lies in striking a balance between engaging tasks and the precise measurement of response latency offered by digital implementations. Recognizing this, recent research has started to gravitate towards digitalizing tasks, leveraging tablets as a platform for more engaging tasks (Willoughby & Blair, 2016). Yet, some studies, such as that by Verhagen et al. (2017), continued to exclusively focus on response accuracy despite the opportunity for integrating response latency measurements provided by the digital platform.

**Challenges in Relying Solely on Response Accuracy**

The rapid cognitive development observed during early childhood emphasizes the importance of incorporating response latency in assessments of inhibitory control. Even as infants, children demonstrate rudimentary signs of inhibitory control, which intensify and refine over time, particularly between the ages of three and six years (Diamond, 2006; Roebers, 2017). As children age, their command of inhibitory control progressively enhances, a progression that becomes visible through their performance in inhibitory tasks.

Children under five, although capable of understanding the demands of the tasks, typically grapple with disregarding irrelevant information and refraining from improper responses (Tamm et al., 2002). In contrast, children aged five and above often display proficiency in such tasks, consistently achieving high response accuracies (Magnus et al., 2019; Roebers, 2017). This consistent high performance gives rise to a ceiling effect. The ceiling effect describes a result pattern in which the majority of the data clusters at the upper limit of the measurement scale, leaving little to no room for scores to increase (Vogt, 2005, p. 40). This scenario is often observed in simple inhibitory tasks where older children can easily master the task, thus resulting in high scores for the group overall. The risk associated with the ceiling effect is its potential to obscure individual differences among older children who consistently exhibit high task accuracy but may have varying levels of inhibitory control. This effect limits the discriminatory power of the task when only investigating response accuracy, making it challenging to distinguish higher performers from their peers, thereby undermining the sensitivity and validity of the measure.

The rapid evolution of inhibitory control during early childhood presents unique challenges in its assessment. In a meta-analysis of 198 studies including inhibitory control tasks, Petersen et al. (2016) found that tests designed to assess inhibitory control in children typically yield informative response accuracy data within a relatively narrow age band. For most tests, this is approximately three years. For instance, the Shape Stroop Task (Kochanska et al., 1997) can accurately measure inhibitory control in children as young as 1.5 years, but its validity wanes after the child reaches 3.5 years. On the other hand, the Simon Says Task (Strommen, 1973) preserves its validity up to approximately age seven but only offers valuable inhibitory control accuracy data from the age of about 4.5 years. Consequently, the appropriateness of the tests for the specific sample under examination becomes crucial when only response accuracy is considered.

**Understanding the Complications of Response Latency**

While accuracy offers an essential measurement of inhibitory control, we must also consider the inclusion of response latency as a supplemental metric, particularly as children mature and display high levels of accuracy. The response latency becomes especially relevant when we observe a ceiling effect in response accuracy; once a significant portion of children consistently achieves high accuracy, the measure may lose its power to discern individual differences. Hence, assessing the response latency offers another dimension of variability, which helps further understand the child's inhibitory control capabilities (Jones et al., 2003).

However, this approach warrants a careful interpretation. The speed-accuracy tradeoff – a well-established cognitive principle that suggests an inverse relationship between the speed of a response and its accuracy (Heitz, 2014) – becomes particularly relevant. Fast but erroneous responses may not necessarily serve as reliable indicators of inhibitory control. Per this principle, rapidly executed responses may lack precision, whereas slower, more deliberate responses often prove more accurate. In the context of inhibitory control, the speed-accuracy tradeoff implies that a child who responds quickly but inaccurately might not possess the same level of inhibitory control as a child who responds a bit slower but accurately. Inhibitory control, specifically response inhibition, refers to the ability to suppress a prepotent or dominant response (Verbruggen & Logan, 2008). In this context, it becomes clear that a fast response in alignment with the prepotent response (i.e., a wrong response) cannot indicate effective inhibitory control. Paap (2019) argues that accuracy might be a more valid metric for young children, who are still developing their skills and tend to be less accurate. Consequently, when we evaluate response latency as a measure of inhibitory control, we need to ensure the measurement is weighed against the accuracy of the response to capture a holistic picture of the child's inhibitory control capabilities.

Therefore, while response latency indeed provides valuable information, particularly in circumstances where response accuracy reaches a ceiling, it also brings its own set of challenges. A comprehensive understanding of inhibitory control should consider both accuracy and response latency, acknowledging the interaction between speed and accuracy as per the speed-accuracy tradeoff. The decision to include response latency as a measurement should be made with the complexity of the task and the proficiency level of the children being tested in mind.

**Utilizing Different Study Designs to Understand Inhibitory Control**

Until now, it appears optimal to prioritize response accuracy as long as it provides adequate discriminatory value, and to consider integrating response latency only when a ceiling effect is observed within the sample. However, diversity in sample characteristics, such as variation in age, developmental stages, and individual abilities, introduces another layer of complexity to the assessment of inhibitory control. In such cases, a more nuanced approach may be required. Understanding and assessing inhibitory control, a complex cognitive ability, poses methodological challenges. Both longitudinal and cross-sectional studies offer valuable insights into children's development and manifestation of inhibitory control. Longitudinal studies examine changes over time, providing stronger evidence for causal relationships and temporal precedence (Duckworth et al., 2010; Farrington, 1991), while cross-sectional studies are relatively more straightforward to conduct and useful for estimating prevalence and examining relationships at a single point in time (Farrington, 1991). In both longitudinal and cross-sectional designs, the children's abilities can vary at any given time, influenced by individual skills or external factors like socioeconomic status (Kałamała et al., 2020; Lipina et al., 2013). In longitudinal studies, the progression of children's development over time leads to significant diversity in their abilities. Such diversity in a sample can lead to scenarios where some children reach a ceiling in response accuracy, making response latency a relevant metric. In contrast, other children may still struggle with accuracy, rendering a focus on response latency premature. This variation in abilities across the sample presents methodological challenges. Yet, it remains crucial for investigating diverse populations, ensuring the generalizability of research findings, and comprehending the developmental trajectory of inhibitory control.

Given the complexity of inhibitory control and the importance of including diverse samples, developing, and utilizing measurement tools that can effectively assess this ability across various populations is essential. By using reliable and valid measures of inhibitory control, researchers can better understand the mechanisms underlying this complex cognitive ability and inform interventions and educational practices to promote its development in children.

**The Combined Approach: Integrating Accuracy and Latency in Measurement**

We encounter a few challenges in understanding the complexity of inhibitory control in children. First, older children often reach ceiling effects in inhibitory tasks (Roebers, 2017), rendering response accuracy an insufficient measure due to the limited range of scores. The consideration of response latency as an alternative might help mitigate this issue. However, response latency is influenced by the speed-accuracy tradeoff (Heitz, 2014), where faster responses may not always correspond to improved inhibitory control, particularly in young children for whom accuracy may be a more valid measure of inhibitory control (Paap, 2019). Thus, the challenge lies in effectively integrating both response accuracy and latency to assess inhibitory control across different developmental stages accurately. Recognizing this necessity, researchers have begun to explore and develop approaches to integrate these two metrics, thereby providing a more holistic and accurate measure of inhibitory control. These approaches may improve the comprehension and assessment of inhibitory control in children.

***Composite Measures: Formula-Based Approaches to Efficiency***

A potential strategy involves employing established formula-based methods to generate composite indices of performance efficiency, effectively integrating the two distinct variables into a cohesive and interpretable score (Liesefeld & Janczyk, 2019; Vandierendonck, 2017). Building on this notion, several researchers have formulated composite measures, each featuring a distinct methodology for encapsulating performance efficiency in a single score. Two of these formula-based approaches are worth highlighting.

The Inverse Efficiency Score (IES; Formula 1), proposed by Townsend and Ashby (1983), integrates response latency and accuracy by dividing the average response latency ($RL$) by the mean accuracy ($1 - PE$; with PE standing for the proportion of errors):

$$IES = \frac{RL}{1 - PE} \tag{1}$$

By focusing on the inverse relationship between response accuracy and latency, the IES provides a composite score for a balanced consideration of speed and accuracy.

As a more sophisticated measure, the Bin-Score (Formula 2), described by Hughes et al. (2014), uses a set of steps to combine response latency and accuracy into one overall performance score. The initial step of this process is the calculation of the average response latency, denoted as $RLc$, over all participants and trials within a control condition. Once $RLc$ has been computed, response latencies from correct trials within the experimental condition are adjusted by subtracting $RLc$, resulting in a collection of adjusted response latencies. This collection is then sorted from smallest to largest and divided into ten deciles or bins.

Each of these bins is then assigned a weight based on its sequential position within the decile range, with bin 1 receiving the lowest weight and bin 10 the highest. Furthermore, a unique 'bad' bin accounts for error trials. The weight or penalty associated with this 'bad' bin is substantially larger than any of the other bins, set at 20. The individual scores for each participant are subsequently calculated by multiplying the number ($n$) of correct adjusted response latencies in each bin ($n_i$) by their respective weights ($i$) and adding the weighted count of error trials from the 'bad' bin ($e$). Mathematically, this scoring is represented by the formula:

$$\left(\sum_{i=1}^{10} n_i * i\right) + n_e * 20 \tag{2}$$

The derived score represents the magnitude of the difference in performance between the experimental and control conditions. It integrates both the accuracy and latency of responses, offering a nuanced performance measure. Lower Bin-Scores are indicative of superior performance, as they represent a higher number of accurate responses and fewer errors in lower latency bins.

The application of these formulae brings several strengths. Foremost among these is their relative simplicity, which facilitates straightforward calculation. Additionally, they produce integrated scores that are comparable across different studies, enhancing the robustness and generalizability of research findings. However, these formula-based approaches also have their limitations. Notably, in developmental psychology, these formulae rest on the assumption that response accuracy and latency information can be interpreted equivalently across all participants. This assumption might not hold true given the significant variation in response latencies and error rates across different ages, developmental stages, and skill levels. Furthermore, while these formulae offer a concise performance summary, they might mask some of the nuanced information contained within separate measures of speed and accuracy. For instance, a participant may exhibit slow but highly accurate responses, or vice versa, and these differences may be critical to understanding the intricacies of inhibitory control.

Therefore, while using formula-based approaches offers practical solutions to the challenges of integrating response accuracy and latency, it is crucial to interpret these composite measures with an understanding of their assumptions and potential limitations. The appropriateness of their use may depend on the study sample's specifics and the investigated sample's cognitive ability. Given these constraints, a more sophisticated and nuanced analysis of these metrics is required. With the development of novel methodologies, such as model-based approaches, we now have the tools to meet this demand.

### *Unveiling Cognitive Dynamics: The Model-Based Approach*

As the name suggests, model-based strategies apply statistical models to explore the multifaceted nature of cognitive processes like inhibitory control. An example of a model-based approach to investigate inhibitory control was proposed by Magnus et al. (2019). They integrated response accuracy and latency into a structural equation model using item response accuracies and latencies as indicators of inhibitory control. The researchers proposed a bifactor model comprising an Inhibitory Control factor based on item response accuracies and latencies, and an orthogonal Response Time factor only based on item response latencies. This Response Time factor aimed to capture general response speed unrelated to inhibitory control. By setting the factors as orthogonal in their model, the authors could distinguish between the variance in response latency indicative of inhibitory control ability and the variance representing general processing speed. The results indicated that inhibitory control scores based on the bifactor model were more precise than those generated from the unidimensional model. Further advancements in this model-based approach were presented by Camerota et al. (2020). They expanded the previous method by incorporating response latencies from a base condition (not requiring inhibition) as indicators for the Response Time factor. This extension allowed for additional control of general processing speed, independent of inhibitory processes.

Such model-based approaches offer the option of exploring cognitive dynamics. They hold the potential to inform our understanding of inhibitory control, revealing the intricate interplay between accuracy and latency. Despite being relatively new, these methods can enable a more refined assessment of cognitive performance, contributing to the continuous evolution of research in developmental psychology and beyond. However, as with any methodology, the successful application of model-based approaches relies on carefully interpreting the models and considering their assumptions and potential limitations. For example, the restriction remains in their limited ability to account for the variability and individual differences within samples, particularly those with a broad age range or ability spectrum. Therefore, while model-based approaches provide valuable insights, they may not be suitable for capturing the full complexity of inhibitory control in diverse populations. The non-linear development of inhibitory control emphasizes this limitation.

**Developmental Complexity: The Non-linear Maturation of Inhibitory Control**

The development of inhibitory control is a complex and non-linear process characterized by both quantitative and qualitative changes over time (Carlson & Moses, 2001; Roebers, 2017). One of the distinctive features of this development is that the relationship between response latency and inhibitory control abilities appears to be contingent on the level of response accuracy. This interaction has been demonstrated in a study by Camerota et al. (2020). The authors could show that longer response latencies were associated with better executive function abilities for children who exhibit lower response accuracy. Conversely, shorter response latencies indicate higher executive function abilities in children who demonstrate high response accuracy. This pattern reflects the critical role of developmental changes in the relationship between response latencies and accuracy, especially in inhibitory control tasks. Neurological research further substantiates this non-linear development in inhibitory control. Younger children show a more generalized neural activation pattern during inhibitory control tasks (Tamm et al., 2002). As children mature, their brain activation becomes more focused, particularly in the left inferior frontal gyrus, a region associated with inhibitory control (Swick et al., 2008). This development underscores a qualitative shift in children's strategies as they become more proficient in inhibitory control tasks.

Considering this intricate relationship between response latency, accuracy, and inhibitory control development, it may be crucial to prioritize accuracy as a key measure in assessing younger children's inhibitory control abilities (Paap, 2019). In children with lower accuracy, shorter response latencies can hardly be seen as an indicator of inhibitory control. They could be interpreted as the inability to suppress the prepotent response long enough to provide the desired answer. Conversely, shorter latencies may suggest more efficient inhibitory control in children with higher accuracy. This nonlinearity of inhibitory control development regarding response latency implies a change in how children manage inhibitory tasks as they age. Slower response latencies may signal a beneficial strategy in younger children or those with lower accuracy, as they take the necessary time to inhibit automatic responses. An analytical approach that universally favors quick response latencies could be inadequate in fully capturing these dynamics.

Two established frameworks support this notion: The Dual Processing Theory (Schneider & Fisk, 1983) and the Horse Race Model (Logan & Cowan, 1984). The Dual Processing Theory (Schneider & Fisk, 1983) differentiates between controlled and automated processing. Longer response latencies, in a context requiring a high degree of control, can indicate an effective, albeit slower, solution process. Conversely, when a task can be completed automatically, shorter response latencies signify an efficient solution process. However, the dynamics between response latency and accuracy in inhibitory control tasks are unique. They are specifically designed to challenge automatic responses and demand conscious regulation, which can result in longer response latencies even when the solution process is effective. While these tasks cannot be automated in the same manner as other cognitive tasks, it can be argued that if children have acquired a more specialized ability for controlling their responses, the potential benefit derived from taking additional time to respond may be reduced. In the context of inhibitory control tasks, the Dual Processing Theory implies a distinction between effective and efficient processing. Longer response latencies in situations requiring a high degree of effort can indicate an effective, albeit slower, solution process. Conversely, shorter response latencies in situations that can be completed more with less effort signify an efficient solution process.

The Horse Race Model is a theoretical framework proposed by Logan and Cowan (1984). It aims to explain inhibitory control processes in response inhibition tasks. It posits the existence of two competing processes: the response time process and the stop-signal process. The response time process is responsible for executing the response. It involves the cognitive and motor processes that allow an individual to select and execute a response based on the presented stimulus quickly and accurately. On the other hand, the stop-signal process is responsible for inhibiting the response. It involves the cognitive processes that enable an individual to suppress or override the prepotent response tendency. This specificity is unique to the stop-signal paradigm. In this paradigm, participants are typically required to respond quickly to a go signal (e.g., pressing a button) but occasionally receive a stop signal (e.g., an auditory cue) that indicates they should withhold their response (Verbruggen & Logan, 2008). In these cases, two explicit signals are present: The reaction signal and the stop signal.

While this specificity differs from tasks like the Stroop Task, the Horse Race Model can be applied to gain insights into inhibitory control in younger children. When participants are required to inhibit their automated responses and to not respond with a prepotent response, there are also two processes at work: The process of answering the item as quickly as possible, which in some cases may be explicitly instructed, and the process of adhering to the rule contradicting the automated response. By applying the principles of the Horse Race Model to a Stroop Task, we can examine the relationship between response latencies and accuracy as indicators of inhibitory control. It suggests a potential difficulty in activating the inhibitory process promptly or effectively enough to halt the instinctual response. In this scenario, the response time process "wins" the race by finishing first, resulting in premature response execution. The fast response latencies observed in these cases may indicate a failure to inhibit the prepotent response fast and long enough to select the correct response rather than reflecting efficient inhibitory control. Only fast response latencies combined with maintained accuracy indicate successful inhibitory control, implying that children can successfully activate their inhibitory control promptly and sustain it long enough to suppress the prepotent response, thereby choosing the correct one. By applying the principles of the Horse Race Model, we can better interpret the interplay between response latencies and accuracy in response inhibition tasks. Overall, short response latencies in children with low accuracy might not indicate efficient inhibitory control but instead suggest a difficulty in suppressing instinctual responses quickly and long enough to answer correctly. Conversely, faster responses may be a testament to efficient inhibitory control for children with high accuracy.

To fully understand the complexity of inhibitory control development and its relationship with response accuracy and latency, it is essential to employ a valid instrument to capture these nuances. Additionally, a diverse sample in terms of age and skill is necessary to ensure the generalizability of findings. A longitudinal design would allow for both cross-sectional and longitudinal evaluations, shedding light on the changes and stability of inhibitory control over time.

**References**

Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Developmental Psychology, 50*(10), 2368–2379. https://doi.org/10.1037/a0037493

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology, 5*(2), 115–129. https://doi.org/10.1076/chin.5.2.115.3167

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8*(4), 485–493. https://doi.org/10.1037/0894-4105.8.4.485

Berger, A., Jones, L., Rothbart, M. K., & Posner, M. I. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. https://doi.org/10.3758/BF03207798

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*(1), 48-71. https://doi.org/10.1016/0273-2297(90)90004-N

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly, 24*(3), 337–349. https://doi.org/10.1016/j.ecresq.2009.06.001

Camerota, M., Willoughby, M. T., Magnus, B. E., & Blair, C. B. (2020). Leveraging item accuracy and reaction time to improve measurement of child executive function ability. *Psychological Assessment, 32*(12), 1118–1132. https://doi.org/10.1037/pas0000953

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Boeck, P., & Wilson, M. (Eds.). (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. Springer. https://doi.org/10.1007/978-1-4757-3990-9

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157-173. https://doi.org/10.1016/0160-2896(91)90028-C

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169539.003.0006

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "do as I say, not as I do". *Developmental Psychobiology, 29*(4), 315–334. https://doi.org/10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO;2-T

Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year-old children. *Psychological Test and Assessment Modeling, 52*(4), 419–431.

Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. *Social Psychological and Personality Science, 1*(4), 311–317. https://doi.org/10.1177/1948550609359707

Farrington D. P. (1991). Longitudinal research strategies: Advantages, problems, and prospects. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(3), 369–374. https://doi.org/10.1097/00004583-199105000-00003

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480

Groppe, K., & Elsner, B. (2014). Executive function and food approach behavior in middle childhood. *Frontiers in Psychology, 5*, 447. https://doi.org/10.3389/fpsyg.2014.00447

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. *Psychology and Aging, 7*(1), 56–64. https://doi.org/10.1037/0882-7974.7.1.56

Hao, J. (2017). Do children with better inhibitory control donate more? Differentiating between early and middle childhood and cool and hot inhibitory control. *Frontiers in Psychology, 8*, 2182. https://doi.org/10.3389/fpsyg.2017.02182

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). Academic Press. https://doi.org/10.1016/B978-012208930-5/50007-6

Harnishfeger, K. K., & Bjorklund, D. F. (1993). The ontogeny of inhibition mechanisms: A renewed approach to cognitive development. In M. L. Howe & R. Pasnak (Eds.), *Emerging themes in cognitive development* (pp. 28-49). Springer. https://doi.org/10.1007/978-1-4613-9220-0_2

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*, 150. https://doi.org/10.3389/fnins.2014.00150

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*(2), 107–129. https://doi.org/10.1016/0001-6918(67)90011-X

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*(3), 702–721. https://doi.org/10.3758/s13428-013-0411-5

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498–504. https://doi.org/10.1111/1467-7687.00307

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition, 204*, 104373. https://doi.org/10.1016/j.cognition.2020.104373

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263–277. https://doi.org/10.2307/1131849

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Lipina, S., Segretin, S., Hermida, J., Prats, L., Fracchia, C., Camelo, J. L., & Colombo, J. (2013). Linking childhood poverty and cognition: Environmental mediators of non-verbal executive control in an Argentine sample. *Developmental Science, 16*(5), 697–707. https://doi.org/10.1111/desc.12080

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*(7), 1296–1306. https://doi.org/10.1177/1073191117740953

Mamrot, P., & Hanć, T. (2019). The association of the executive functions with overweight and obesity indicators in children and adolescents: A literature review. *Neuroscience and Biobehavioral Reviews, 107*, 59–68. https://doi.org/10.1016/j.neubiorev.2019.08.021

Moriguchi, Y., & Hiraki, K. (2013). Prefrontal cortex and executive function in young children: A review of NIRS studies. *Frontiers in Human Neuroscience, 7*, 867. https://doi.org/10.3389/fnhum.2013.00867

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*(2), 111–121. https://doi.org/10.1111/j.2044-8295.1990.tb02349.x

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*(4907), 933–938. https://doi.org/10.1126/science.2658056

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

Monsell S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134–140. https://doi.org/10.1016/s1364-6613(03)00028-7

Müller, U., & Kerns, K. (2015). The development of executive function. In L. S. Liben, U. Müller, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (pp. 571–623). John Wiley & Sons. https://doi.org/10.1002/9781118963418.childpsy214

Ng, F. F.-Y., Tamis-LeMonda, C. S., Yoshikawa, H., & Sze, I. N.-L. (2014). Inhibitory control in preschool predicts early math skills in first grade: Evidence from an ethnically diverse sample. *International Journal of Behavioral Development, 39*(2), 139–149. https://doi.org/10.1177/0165025414538558

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Paap, K. R. (2019). The bilingual advantage debate: Quantity and quality of the evidence. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch34

Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., & Staples, A. D. (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review, 40*, 25–71. https://doi.org/10.1016/j.dr.2016.02.001

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social–emotional competence. *Journal of Applied Developmental Psychology, 30*(3), 310–320. https://doi.org/10.1016/j.appdev.2008.12.012

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Schneider, W., & Fisk, A. D. (1983). Attentional theory and mechanisms for skilled performance. In R. A. Magill (Ed.), *Memory and control of action* (pp. 119–143). North-Holland. https://doi.org/10.1016/ S0166-4115(08)61989-5

Simpson, A., & Carroll, D. J. (2019). Understanding early inhibitory development: Distinguishing two ways that children use inhibitory control. *Child Development, 90*(5), 1459–1473. https://doi.org/10.1111/cdev.13283

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*(1), 5–42. https://doi.org/10.1006/cogp.1997.0658

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*(5408), 1657–1661. https://doi.org/10.1126/science.283.5408.1657

Strommen, E. A. (1973). Verbal self-regulation in a children's game: Impulsive errors on "Simon says". *Child Development, 44*(4), 849–853. https://doi.org/10.2307/1127737

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience, 9*, 102. https://doi.org/10.1186/1471-2202-9-102

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 37*(4), 571–590. https://doi.org/10.1080/14640748508400920

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(10), 1231–1238. https://doi.org/10.1097/00004583-200210000-00013

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press.

Traverso, L., Viterbori, P., & Usai, M. C. (2015). Improving executive function in childhood: Evaluation of a training intervention for 5-year-old children. *Frontiers in Psychology, 6*, 525. https://doi.org/10.3389/fpsyg.2015.00525

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Verhagen, J., Mulder, H., & Leseman, P. P. M. (2017). Effects of home language environment on inhibitory control in bilingual three-year-old children. *Bilingualism: Language and Cognition, 20*(1), 114–127. https://doi.org/10.1017/S1366728915000590

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Vogt, W. P. (2005). *Dictionary of statistics and methodology* (3rd ed.). Sage.

Willoughby, M. T., & Blair, C. B. (2016). Longitudinal measurement of executive function in preschoolers. In J. A. Griffin, P. McCardle, & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 91–113). American Psychological Association. https://doi.org/10.1037/14797-005

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2011). Contributions of modern measurement theory to measuring executive function in early childhood: An empirical demonstration. *Journal of Experimental Child Psychology, 108*(3), 414–435. https://doi.org/10.1016/j.jecp.2010.04.007

Yücel, M., Fornito, A., Youssef, G., Dwyer, D., Whittle, S., Wood, S. J., Lubman, D. I., Simmons, J., Pantelis, C., & Allen, N. B. (2012). Inhibitory control in young adolescents: The role of sex, intelligence, and temperament. *Neuropsychology, 26*(3), 347-356. https://doi.org/10.1037/a0027693

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 445–469). Blackwell Publishing. https://doi.org/10.1002/9780470996652.ch20

# Chapter II

# Present Research

# CROCODILE Project

In this section, I will provide an overview of the CROCODILE project. This project provides the context for the research work carried out and presented in this dissertation.

## Summary of the CROCODILE Project

The CROCODILE project, short for **CRO**ss-national Interdis**C**iplinary Study **O**n Child **D**evelopment **I**n **L**inguistically-diverse **E**nvironments, began in 2019 as a longitudinal study conducted over four years. Over this time, it investigated the linguistic, socio-emotional, metacognitive, and cognitive development of single and dual language learners aged 3 to 6 across Switzerland and Germany. This project was the result of an interdisciplinary collaboration between four universities: the University of Basel, the University of Bern, and the University of Neuchâtel in Switzerland, as well as the University of Würzburg in Germany.

Throughout the course of the study, we aimed to assess each participating child three times, with time intervals ranging between 6 to 12 months. This repeated assessment approach allowed us to effectively track the development of various skills in the children over the span of up to two years. The CROCODILE project aimed to include children with diverse language backgrounds. Participants were required to have one of the following language combinations: single language learners of (Swiss) German or French or dual language learners with the societal language of (Swiss) German or French and the heritage language of Italian or Turkish.

## Study Design

The CROCODILE project featured three measurement points, which formed the longitudinal study. In addition to the three main measurement points, two piloting phases were conducted to ensure the effectiveness and suitability of the tests used in the project. The first piloting phase occurred before the start of the longitudinal study. This preliminary assessment served to verify that the chosen tests effectively captured the intended developmental aspects in the children while also providing the opportunity to make any necessary adaptations. The second piloting phase took place between the first and second measurement points. This phase was dedicated to assessing the performance of newly introduced tests.

Each study session was tailored to the children's language setting. Single language learners were assessed in their respective language, while for dual language learners, parents were asked to identify their child's more proficient language. This language was then used as the test language, with the other language being the non-test language. All instructions were given in the test language, except for assessments specifically designed to measure both languages, such as vocabulary tests. In these instances, the children were required to undergo the assessment twice: initially in their designated test language and subsequently in their non-test language. The instructions for each assessment iteration were presented in the corresponding language of that particular evaluation round. If a research assistant noticed that a child was more proficient in the non-test language during the assessment, the experiment's language would be switched accordingly.

The duration of the assessments varied across the different stages of the project. The first piloting phase consisted of two 90-minute sessions, while the second piloting phase took one 90-minute session. The first timepoint of the longitudinal study consisted of two 90-minute sessions, whereas the second and third waves each comprised a single 120-minute session. During the study, most assessments were computerized and guided by a virtual crocodile character named Sammy (Figure 1: Panel A, C). Dual language learners were also introduced to Lilly (Figure 1; Panel B, D), another crocodile who spoke the non-test language. The children accompanied these characters on a virtual treasure hunt or helped them find their way home in the various tests, serving as a consistent and engaging cover story. In non-computerized tests, the research assistant used a hand puppet version of Sammy to maintain continuity. At the end of each measurement point, the children received a small gift as a token of appreciation for their participation in the study.

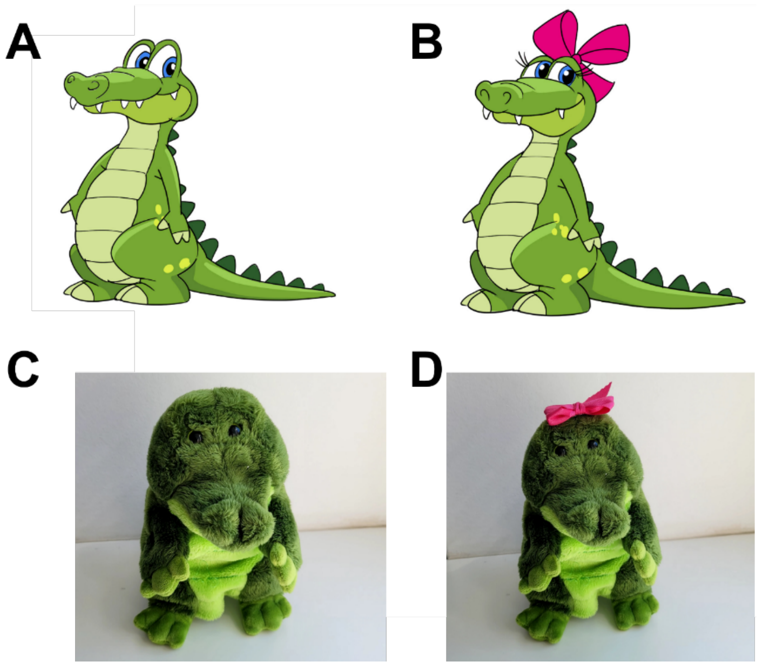
**Inhibitory Control Assessment**

The computerized Pointing Stroop Task (cPST) was a core component of the test battery used in the CROCODILE project. Based on the work of Berger et al. (2000), a modified version was implemented to suit the characteristics and needs of the diverse sample of young children in this study. The cPST was administered at every timepoint, providing a consistent measure of inhibitory control throughout the study.

The cPST engaged the children in a story where they, along with the character Sammy, encountered a dolphin trying to learn the sounds made by a dog and a cat (Figure 2, Panel A). The task was divided into two blocks, each consisting of eight trials in the longitudinal studies and four trials in the piloting study. It took about five minutes to complete, including the story, instructions, practice trials, and the test itself. In both blocks, children were shown images of a cat and a dog on the screen, with their left and right positions being randomized (Figure 2, Panel B). Simultaneously, they heard either a *bark* or a *meow* sound. In the first block, the congruent block, the children were asked to press the image matching the animal sound they heard. This instruction meant pressing the image of the cat when they heard a *meow* and the image of the dog when they heard a *bark*. This block required a straightforward association of sound to image. The second block, the incongruent block, challenged the children's inhibition. In this block, children were instructed to press the image contradicting the animal sound, meaning they had to press the image of the dog when hearing a *meow* and the image of the cat when hearing a *bark*. This block required inhibitory control, as the children had to suppress their instinct to associate the sound with the corresponding image and instead do the opposite.

**Figure 1**

*Digital Versions and Puppets of the Characters "Sammy" and "Lilly"*




*Note.* Digital (Panel A, B) and Puppet (C, D) Versions of Sammy (A, C) and Lilly (B, D). They were used as (virtual) characters to guide the children through the experiment and ensure continuity within and between measurement points.

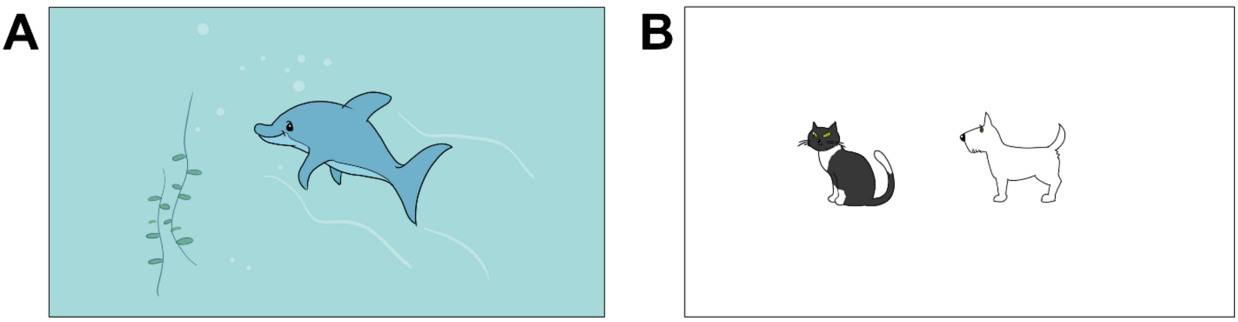
Before starting each block, children had to correctly solve two practice trials to familiarize themselves with the task. If they made a mistake, the practice trial was repeated until they could perform it correctly. The task was designed to be engaging, fun, and challenging, yet appropriate for the age group, and aimed at measuring inhibitory control accurately. To ensure that the task was fitting for our diverse sample of young children in terms of age, language, and cultural background, instructions were presented in the language the child was most proficient in. This approach ensured that all children, regardless of their language and cultural background, had an equal opportunity to understand and engage in the task, eliminating any language-related biases. Additionally, the necessity of completing both practice trials correctly further guaranteed that children understood the instructions. These methodological considerations were crucial for investigating whether inhibitory control can be accurately measured across a broad age and skill range through appropriate assessment methods, a central research question of this dissertation.

**Overall Sample**

The sample size varied across different measurement points of the CROCODILE project. A detailed representation of the sample distribution for children completing the computerized Pointing Stroop Task is provided in Table 1. In sum, the study encompassed 526 unique children who participated in the computerized Pointing Stroop Task. This diverse and longitudinal sample ensured a robust examination of inhibitory control across different developmental stages.

**Figure 2**

*Screenshots from the Story and Instruction and the Trials of the computerized Pointing Stroop Task*




*Note.* Panel A: One of the images presented to the children during the story and explanation of the task. Panel B: The clickable images of the cat and the dog during the (practice) trials.

**Table 1**

*Sample Distribution Across Measurement Points*

| Measurement Point | *n* total | *n* bilinguals | *n* monolinguals |
|---|---|---|---|
| Piloting 01 | 135 | 43 | 92 |
| Longitudinal 01 | 322 | 125 | 197 |
| Longitudinal 02 | 355 | 130 | 225 |
| Longitudinal 03 | 281 | 94 | 187 |

*Note.* Number of participants that took part in the study and completed the computerized Pointing Stroop Task across the measurement points.

## Overview, Motivation, and Aim of the Present Work

This dissertation explores the challenges related to the rapid and non-linear development of inhibitory control in children. As recent scientific trends incorporate response accuracy and latency in evaluating inhibitory control (e.g., Camerota et al., 2019; Camerota et al., 2020; Magnus et al., 2019; Zelazo et al., 2013), this dissertation seeks to contribute significantly to this emerging paradigm. My objective is to enhance the precision and validity of measurement methodologies, thereby gaining deeper insights into the maturation of inhibitory control in children and identifying more reliable assessment methods to understand its intricate development. With that foundation, I aim to investigate inhibitory control in young children by considering response accuracy and response latency in tandem, to gain new insights into its development.

My starting point is to confirm the validity of the research instrument employed. Chapter 3 presents the first study, where I evaluate the validity of response accuracy and latency in a basic inhibitory control test – the computerized Pointing Stroop Task (cPST; Berger et al., 2000). Moreover, this investigation explores the association between performance efficiency, a combination of speed and accuracy, and fluid intelligence. Chapter 4 presents the second empirical study, addressing the challenges in assessing inhibitory control due to its rapid (Kochanska et al., 1996) and non-linear development (Camerota et al., 2020). I use the cPST validated in study 1 to analyze the relationship between response accuracy, response latency, and inhibitory control, using a cross-sectional approach while making distinctions between younger and older children. The intent is to understand how inhibitory control manifests differently across varying developmental stages and skill levels. This chapter further evaluates the possibility of combining response accuracy and latency into a singular metric. As the final step Chapter 5 explores the longitudinal development of the relationship between response latency, accuracy, and inhibitory control. This perspective helps to understand the dynamic relationship among these variables across time and developmental stages. This chapter investigates whether response latency can be a credible marker of inhibitory control and probes the extent to which this association hinges upon accuracy levels.

These empirical chapters provide a thorough investigation, untangling the intricate connections between inhibitory control, response accuracy, and response latency, contributing to a robust understanding of these crucial components of cognitive development and potentially facilitating future valid evaluations of children's cognitive capabilities. Finally, the potential impact of this work extends beyond academia, having implications for educational practices and psychological interventions. The findings can help develop pedagogical strategies and psychological interventions that duly consider the intricacies of inhibitory control development in children.

## Theoretical Motivation

The complexity of cognitive development, particularly inhibitory control, requires nuanced assessment methods to enable researchers to track the development across a broad age range. Be it longitudinally or cross-sectionally. Inhibitory control not only occupies a pivotal role in the cognitive development of children but also intertwines with broader cognitive capacities, such as fluid intelligence (Dempster, 1991; Roca et al., 2010), highlighting its importance in the overall cognitive maturation process. This connection underlines the importance of assessing inhibitory control precisely and validly. Traditional approaches to assess inhibitory control are hampered by the rapid developmental pace of children (Kochanska et al., 1996), often leading to a ceiling effect in response accuracy measurements (Roebers, 2017). Recent research has underscored the necessity of incorporating both response accuracy and latency in evaluating inhibitory control (Camerota et al., 2019; Magnus et al., 2019). However, this poses challenges, especially considering the non-linear and complex development of response latency across different childhood stages (Camerota et al., 2020). This dissertation, therefore, explores the premise that a nuanced integration of response accuracy and latency while acknowledging their dynamic relationship can offer a refined and comprehensive assessment of inhibitory control.

**Aim of the Present Research**

The overarching goal of this dissertation is to shed light on the developmental complexities of inhibitory control in children and provide methods and starting points to measure inhibitory control more precisely and accurately. My objective is to explore if and how integrating response accuracy and latency can advance this aspect. In pursuing this aim, the dissertation seeks to further examine how these two dimensions of inhibitory control evolve from early childhood through critical developmental stages. Additionally, I explored various analytical approaches to response accuracy and latency in inhibitory control tasks, aiming to identify which derived scores demonstrate a stronger correlation between inhibitory control and fluid intelligence. This approach not only facilitated the examination of various assessment methods' validity but also offered initial insights into uncertainties regarding the relationship between inhibitory control and other cognitive or behavioral aspects of human experience.

Specifically, this dissertation sets out to achieve the following goals: First, I want to lay a foundation for the following investigations by validating the test used in later studies to assess inhibitory control – namely, the computerized Pointing Stroop Task (Berger et al., 2000). Corroborating the validity of that measure both in terms of response accuracy and latency enables us to investigate the following research questions and hypotheses with higher confidence. Second, this dissertation will investigate how the integration of these two metrics can enhance the picture of inhibitory control beyond looking at the metrics in isolation. Third, this work will explore how developmental stages, both in terms of age and proficiency, influence the relationship between response accuracy and latency and what information they provide for the assessment of inhibitory control. Fourth, this study aims to explore different assessment approaches to not only validate the integration of these two metrics but also to provide actual, tangible, practical knowledge both for future research as well as providing starting points for practitioners. Fifth, I am to investigate how the knowledge gained in the process of this dissertation enables a nuanced look and investigation of the development of inhibitory control in young children. Lastly, the study will examine how the assessment methods can influence the investigation of the relationship between inhibitory control and related cognitive concepts, for example, fluid intelligence.

By addressing these goals, the dissertation aims to contribute to the theoretical and empirical advancement of cognitive assessment methods, particularly for young children. It seeks to offer a framework for interpreting inhibitory control metrics, providing critical insight into the developmental transitions from accuracy-dependent to latency-dependent inhibitory control. The results may enable future research to investigate associations of inhibitory control more nuancedly.

## Empirical Investigations in this Dissertation

The following section, I will provide a short exploration of the empirical investigations conducted in this dissertation, each contributing uniquely to the stated research goals.

### *Study 1: Laying the Groundwork*

Selecting an appropriate instrument is essential to assess the development of inhibitory control in young children. Neither response accuracy alone (Roebers, 2017) nor response latency alone (Paap, 2019) comprehensively measures inhibitory control ability across a broad age range. Thus, an ideal task should simultaneously assess response accuracy and precisely measure response latency. Traditional tasks designed for young children, often conducted offline, have limitations in accurately capturing response latency (for a review, see Montgomery & Koeltzow, 2010). In contrast, the computerized Pointing Stroop Task (cPST; Berger et al., 2000) stands out as a suitable instrument for this purpose.

The cPST leverages the congruency effect to operationalize inhibitory control. The task consists of two blocks: the first requires no inhibition, aligning responses with prepotent tendencies, while the second block comprises items that necessitate inhibiting these prepotent responses. This distinction is pivotal, as it marks the primary difference between the items in the two blocks. A valid measurement of inhibitory control using the cPST would be indicated by performance disparities predominantly influenced by the requirement of inhibition in the incongruent block as opposed to the congruent block. Additionally, position effects within the task (e.g., MacLeod & Dunbar, 1988) may also contribute to performance variations, warranting consideration in the task's analysis.

To investigate the validity of the cPST, we applied explanatory item response models (Hartig et al., 2012) to the data collected from the children in the first piloting study. Initially, *empirical item difficulties and time intensities* were estimated through a one-parameter logistic model for response accuracy and a linear mixed model for response latency. Subsequently, *estimated item difficulties and time intensities* were calculated by modeling the performance metric as a function of the item's congruency (whether it was congruent or incongruent) and its presentation position within the test block. Finally, correlations between empirical and estimated item difficulties and time intensities were examined to assess the amount to which participant's performance in the cPST was influenced by item position and, crucially, congruency. Secondly, we explored the integration of response accuracy and latency to determine whether combining these metrics provides additional insights beyond their isolated analysis. The focus was on understanding their relationship with intelligence. We conducted a moderated linear regression model (Aiken & West, 1991) using mean response accuracy, mean response latency, and their interaction as predictors. The dependent variable in this model was the children's fluid intelligence scores. This approach allowed us to examine how the interaction between response accuracy and latency could explain variations in fluid intelligence.

These investigations have the potential to validate the cPST as an effective tool for assessing inhibitory control. Moreover, they can offer initial insights into the benefits of integrating response accuracy and latency. This study lays the groundwork for further exploration into a holistic understanding of cognitive development.

***Study 2: Cross-sectional Investigations and Consideration of Developmental Stages***

With an instrument that validly and precisely measures inhibitory control through both response accuracy and latency, and initial evidence supporting the integration of these metrics for enhanced insights, the next challenge is to determine the most effective integration method. Established research indicates that inhibitory control undergoes rapid improvement, often leading to ceiling effects in response accuracy (Kochanska et al., 1996; Roebers, 2017). Moreover, recent studies suggest that while response accuracy offers a consistent interpretation across all ages (higher accuracy indicates better inhibitory control), response latency reveals more nuanced information. Specifically, for younger children, longer response latencies may signify stronger inhibitory control (Camerota et al., 2020). This observation aligns with the fundamental concept of behavioral inhibition, which involves suppressing a rapid, prepotent response in favor of a more deliberate and ultimately correct response (Verbruggen & Logan, 2008).

In a cross-sectional study, we compared a two-factor model that integrated response accuracy and latency against a unidimensional model, focusing solely on accuracy as outlined by Magnus et al. (2019). The participants included children assessed with the cPST at the first study timepoint. We aimed to determine whether incorporating response latency could enhance the precision of inhibitory control measurement. Additionally, we investigated whether response latency offered different insights for younger versus older children within the same model structure. Another key aspect of our analysis was to explore how children's proficiency, as indicated by their response accuracy, influenced the relationship between response latency and inhibitory control ability. The hypothesis that longer response latencies could be beneficial for children who make more errors would gain support if we found that increased latencies led to fewer mistakes. In such cases, extended response latencies might not indicate weaker inhibitory control. Finally, to place our findings within a broader cognitive development context, we analyzed the correlation between inhibitory control metrics – integrating response accuracy and latency – with fluid intelligence. Based on the premise of a non-linear relationship between response latency and inhibitory control ability across various ages and proficiency levels, I expect metrics to consider this phenomenon to model inhibitory control more validly.

The second study has the potential to uncover intricate patterns in the development of inhibitory control, investigating whether younger children might profit from a more deliberate and controlled approach to processing, underscoring the developmental significance of taking time for accurate responses. Conversely, older children may demonstrate more adept handling of the interplay between speed and accuracy, indicative of a more refined and efficient inhibitory control mechanism. This differentiation between age groups can highlight the maturation of cognitive processes and has the potential to provide valuable insights into the developmental trajectory of inhibitory control.

***Study 3: Longitudinal Investigations***

The design and cross-sectional analyses conducted in studies 1 and 2 provide essential frameworks for investigating the development of inhibitory control in young children. These studies investigate the hypothesis that a nuanced integration of response accuracy and response latency offers a more detailed depiction of inhibitory control abilities than either measure alone. The next logical step is to extend this assessment into a longitudinal framework, which is the focus of Study 3. I built upon the insights gained by analyzing the relationship between response accuracy, response latency, and inhibitory control in a cross-sectional context. By employing a longitudinal design, I aimed to investigate the trajectories of these relationships over time. The central question I sought to answer is: How do response latency and accuracy develop in concert across multiple time points, and how do these developmental patterns relate to inhibitory control and fluid intelligence?

Methodologically, Study 3 utilized the longitudinal data from children who participated at multiple time points. To capture the evolution of inhibitory control capabilities, we analyzed the growth curves of response latency over the three timepoints using linear mixed-effects models (Baayen et al., 2008), which can handle the intricate interplay of age, growth rates, and individual differences among children. One particular aim was to examine whether early proficiency in inhibitory control, as marked by a high response accuracy, predicts subsequent development of inhibitory control regarding response latency and fluid intelligence.

From an applied perspective, the findings from Study 3 can contribute to interventions tailored to the developmental needs and potential of individual children. This level of personalized insight can result in more targeted interventions, enabling practitioners to focus on fostering particular aspects of cognitive growth at optimal times. In summary, study 3 aims to trace the development of inhibitory control over time, offering a dynamic perspective on cognitive maturation. By understanding the longitudinal relationships between response latency, accuracy, and broader cognitive development, I expect to gain novel insights into the mechanisms governing cognitive control and their significance in the broader landscape of child development.

## Expected Contributions

This dissertation encapsulates a multi-faceted exploration into the intricate development of inhibitory control, providing both the scientific community and practitioners with valuable insights and indicating potential applications. At its core, the dissertation drives forward a nuanced understanding of cognitive maturation in children.

### *Contributions to the Research Field*

Firstly, this dissertation aims to provide an evidence-based methodology. The integration of response accuracy and latency in assessing inhibitory control offers a verified methodological advancement. By presenting thorough empirical analyses that validate the use of both metrics, the dissertation equips researchers with a robust framework to dissect cognitive inhibitory processes with greater precision. Second, insights gained from this dissertation have the potential to provide insights into the developmental trajectories, especially the non-linear progression of inhibitory control development in children. This longitudinal evidence may enhance the existing body of literature, detailing how cognitive abilities unfold over time and augmenting existing developmental theories. Lastly, the rich empirical evidence laid out across the chapters can serve as a blueprint for future studies examining inhibitory control or other cognitive faculties. It provides a methodological template to assess cognitive processes comprehensively, which could be adapted and applied to various cognitive domains.

***Practical Applications***

Firstly, the validated approach to measure inhibitory control may equip practitioners with reliable tools to evaluate cognitive development in children. This aspect is particularly valuable in educational settings for identifying children who may benefit from targeted cognitive training programs. Further, the nuanced understanding of how inhibitory control matures over time advances the scope for designing interventions. The findings can inform strategies that prioritize accuracy before response latency in cognitive training, thereby optimizing the developmental support provided to children. Lastly, the dissertation establishes a link between early inhibitory control performance and later cognitive development, highlighting the potential for early assessment to predict subsequent cognitive abilities. Practitioners can leverage this insight to identify and support children at risk of cognitive developmental delays.

In summary, while the main takeaways are in the field of research, the dissertation goes beyond the pure empirical realm. It delivers first suggestions for practical, evidence-backed applications that can enhance cognitive assessments and support the development of children.

## References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Berger, A., Jones, L., Rothbart, M. K., & Posner, M. I. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. https://doi.org/10.3758/BF03207798

Camerota, M., Willoughby, M. T., & Blair, C. B. (2019). Speed and accuracy on the Hearts and Flowers task interact to predict child outcomes. *Psychological Assessment, 31*(8), 995–1005. https://doi.org/10.1037/pas0000725

Camerota, M., Willoughby, M. T., Magnus, B. E., & Blair, C. B. (2020). Leveraging item accuracy and reaction time to improve measurement of child executive function ability. *Psychological Assessment, 32*(12), 1118–1132. https://doi.org/10.1037/pas0000953

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157-173. https://doi.org/10.1016/0160-2896(91)90028-C

Hartig, J., Frey, A., Nold, G., & Klieme, E. (2012). An application of explanatory item response modeling for model-based proficiency scaling. *Educational and Psychological Measurement, 72*(4), 665–686. https://doi.org/10.1177/0013164411430707

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490–507. https://doi.org/10.2307/1131828

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 126–135. https://doi.org/10.1037/0278-7393.14.1.126

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*(7), 1296–1306. https://doi.org/10.1177/1073191117740953

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*(3), 308–330. https://doi.org/10.1016/j.dr.2010.07.001

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain: A Journal of Neurology, 133*(1), 234–247. https://doi.org/10.1093/brain/awp269

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

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

# Chapter III

# Study 1

# Using Accuracy and Response Times to Assess Inhibitory Control in Kindergarten Children: An Analysis with Explanatory Item Response Models

A version of this chapter was published in:

Schulz, D., Richter, T., Schindler, J., Lenhard, W. & Mangold, M. (2023). Using accuracy and response times to assess inhibitory control in kindergarten children: An analysis with explanatory item response models. *Journal of Cognition and Development, 24*(1), 82-104. https://doi.org/10.1080/15248372.2022.2119977

Data and R-code for the analyses reported in this chapter are available at the repository of the Open Science Framework (https://osf.io/nckpb/?view_only=f23b504b4db5445488222aa65fe98bec).

**Abstract**

Inhibitory control is a core executive function that develops during childhood and is measured with tasks that require the inhibition of a dominant response. The current study examined the diagnostic value of using response accuracy and latency in a simple inhibitory control test, the computerized Pointing Stroop Task (cPST), for kindergarten children. The cPST was completed by 135 children, ages 3 through 6 years with diverse national and cultural backgrounds. In explanatory response models, item difficulties and time intensities could be predicted very reliably by congruency and item position, with incongruent responses causing more errors and longer response latency. Moreover, the prediction of fluid intelligence (a close correlate of inhibitory control) from children's performance in the cPST was enhanced by using response accuracy and response latency, which had a multiplicative effect, indicating that efficient (accurate and fast) inhibitory control is related to fluid intelligence. These results suggest that measuring the efficiency of inhibitory control in young children is a more appropriate assessment than using either response accuracy or response latency.

*Keywords:* inhibitory control, executive function, response time, accuracy, assessment

## Using Accuracy and Response Times to Assess Inhibitory Control in Kindergarten Children: An Analysis with Explanatory Item Response Models

Inhibitory control, especially the ability to suppress inappropriate reactions and responses (Verbruggen & Logan, 2008), is crucial to human action and cognition (Diamond, 2013). The development of inhibitory control, its antecedents, and consequences are also a major topic in developmental research (Bialystock, 2017; Campbell et al., 2020; Paap, 2019, Roebers, 2017). Despite the importance of the concept, measuring inhibitory control during childhood is a challenge. Over the course of early childhood, inhibitory control manifests itself differently depending on the age of the children (Roebers, 2017). Response accuracy in simple inhibition tasks is widely seen as the defining measure of inhibitory control in very young children (e.g. Paap, 2019). However, this measure no longer suffices on its own to assess inhibitory control in children older than five years (Roebers, 2017). Still, many inhibitory control tests for young children are conducted offline and manually and therefore lack an objective measurement of response latency (e.g., Escobar et al., 2018). Considering response time and response accuracy to measure the efficiency of inhibitory control might be the key to assess inhibitory control in the age range from 3 to 6, the age in which children rapidly improve their inhibitory control ability (Carlson & Moses, 2001; Kochanska et al., 1996). In the following sections, we discuss inhibitory control as one of three executive functions (Diamond, 2013) and its development, especially in the kindergarten age. One important aspect for the present study is the relationship between inhibitory control and fluid intelligence. We also discuss the strengths and weaknesses of commonly used test instruments for assessing inhibitory control in 3- to 6-year-olds and will elaborate on the advantages of using response accuracy and latency measures when assessing the efficiency of inhibitory control. We then present a modified version of an established inhibitory control task, the computerized Pointing-Stroop task (cPST; Berger et al., 2000) that can be used to economically measure efficient inhibitory control in 3- to 6-year-olds. The core of the present study is the evaluation of the validity of this diagnostic approach by using explanatory item-response modeling (De Boeck & Wilson, 2004) for response accuracy and response latency. We further examine the construct validity of the cPST in terms of response accuracy and response latency, particularly its sensitivity for developmental changes and the relationship between the children's performance in the cPST and their performance in a non-verbal fluid intelligence test.

**Inhibitory Control**

Inhibitory control is a central cognitive ability that belongs to the executive functions, alongside working memory and cognitive flexibility (Diamond, 2013) or alongside updating and shifting (Miyake & Friedman, 2012), depending on the theoretical perspective. Among the executive functions, inhibitory control may be regarded as a core factor, together with working memory, although the exact cognitive-functional and developmental relationship between the two types of executive functions is a point of contention (for different conceptions, see Miyake & Friedman, 2012; Munkata et al., 2011; Roberts & Pennington, 1996; for an overview, see Müller & Kerns, 2015).

Inhibitory control is an effortful cognitive activity that can take different forms. According to Nigg's (2000) taxonomy, one may distinguish between interference control (inhibition of irrelevant stimulus information), cognitive inhibition (inhibition of irrelevant information in working memory), behavioral inhibition (inhibition of prepotent responses) and oculomotor inhibition (inhibition of reflexive eye movements). Among these types of inhibitory control, behavioral inhibition (also called response inhibition) is "most straightforwardly associated with active suppression and executive functioning" (Friedman & Miyake, 2004, p. 104) and may be considered "a hallmark of executive control" (Verbruggen & Logan, 2008, p. 418; see also Norman & Shallice, 1986). Behavioral or response inhibition is the ability to suppress inappropriate reactions, thereby facilitating goal-directed behavior that conflicts with automated or learned behavior. It may be distinguished psychometrically from other types of inhibitory control but is strongly related to distracter inhibition (Friedman & Miyake, 2004). Response inhibition tasks may be further classified into “hot” vs. “cold” inhibition tasks, i.e. those that involve a motivational conflict (such as in tasks requiring delay of gratification) or not (Garon, 2016). Finally, Simpson and Carroll (2019) have pointed out that response inhibition tasks sometimes require the suppression of a prepotent response over relatively short periods of time in the range of seconds (e.g., Stroop-like tasks) or over longer periods of time in the range of minutes or even longer (e.g., delay-of-gratification tasks). In this study, we focus on the variant of response inhibition that may be classified as "cold" and requires the suppression of a prepotent response over relatively short periods of time.

The suppression of prepotent responses incurs cognitive costs that may manifest in increased error rates and response latencies when confronted with distracting, conflicting, or irrelevant stimuli (e.g., MacLeod, 1991). Conversely, more accurate and faster responses in tasks that involve the suppression of inappropriate reactions are indicative of a person’s response inhibition ability. Executive functions are central to the human cognitive system (e.g., Best et al., 2011). Inhibitory control, in particular, plays a role in different areas of human cognition such as the false-belief understanding aspect of the theory of mind (Devine & Hughes, 2014) or academic success (Ng et al., 2015).

**Inhibitory Control and Fluid Intelligence**

One major correlate of inhibitory control is fluid intelligence, defined as the ability for reasoning, abstract thinking, and problem solving, for example, in the perception of relations between different stimuli (Horn & Cattell, 1967). Theory and empirical research suggest a strong relationship between inhibitory control and fluid intelligence, directly (Diamond, 2013) and in conjunction with the other executive functions (Kane & Engle, 2002). Dempster

(1991) argued that inhibitory control is a central ability in problem solving as it requires the inhibition of task-irrelevant information. The inhibition of task-relevant information also saves working memory resources that can be invested in the processing of task-relevant information and problem solving (Bjorklund & Harshfeger, 1990). Conway et al. (2003) demonstrated that working memory capacity is strongly related to fluid intelligence, and inhibition is assumed to assist this relationship (e.g., Kane & Engle, 2002). Furthermore, inhibitory control is strongly related to the functionality of the prefrontal cortex (PFC) and is perhaps its core function (Dempster, 1991), which has been supported by studies using near infrared spectroscopy (for a review see Moriguchi & Hikari, 2013). Corroborating the relationship between inhibitory control and fluid intelligence, Roca et al. (2010) found that lesions in the PFC negatively correlate with fluid intelligence. Finally, the inhibition of irrelevant information is also important for the disengagement from previously unsuccessful or no longer appropriate solution attempts and thus for exploring alternative solutions (Dempster, 1991).

Despite these arguments for a strong link between inhibitory control and fluid intelligence, studies directly assessing their correlations have produced mixed results. Yücel et al. (2012), for example, found that higher intelligence in adolescents was associated with higher response accuracy in a Stroop task, which is a measure of inhibitory control. Although Friedman et al. (2006) found a positive relationship between inhibitory control and multiple intelligence measures in young adults, further analysis suggested that this relationship was likely due to variance shared by inhibitory control and the updating aspect of executive functions. The relationship with intelligence remained significant for updating but not for inhibitory control when updating and inhibitory control were considered simultaneously as predictors of intelligence. This pattern was later replicated by Duan et al. (2010) for adolescents.

All three research groups used either response accuracy or response latency in the inhibition tasks as the measure of inhibitory control. Using both indicators might provide a more comprehensive picture of inhibitory control abilities because inhibitory control can only be considered efficient to the degree that it is both accurate and rapid (cf. Eysenck & Calvo, 1992). Given the theoretical and neurological links of inhibitory control and fluid intelligence, a positive relationship of efficient response inhibition and fluid intelligence can be expected.

**The Development and Assessment of Inhibitory Control During Early Childhood**

Executive functions improve rapidly during early childhood. Research suggests that executive functions, including inhibitory control, can be measured in children as young as 1 year (Diamond, 2006), but substantial leaps in its development occur throughout the first 6 years of life (Roebers, 2017). Children's performance in inhibition tasks improves rapidly, especially between the ages of 3 and 6, which indicates that preschool children increasingly master inhibitory control (Carlson & Moses, 2001). Most children younger than five years are able to describe the inhibition task they are asked to complete, which demonstrates an understanding of the tasks. Nonetheless they often fail to suppress irrelevant information and inappropriate reactions (Tamm, 2002). Children older than five, however, usually perform well in these tasks, which is reflected in high response accuracies, often to the point of a ceiling effect (Magnus et al., 2019; Roebers, 2017). Consequently, individual differences in inhibitory control seem to be reflected quite well in accuracy measures for younger children. Older children usually show high accuracy in simple inhibition tasks but a large variance in response latencies (Jones et al., 2003).

Notwithstanding this developmental pattern, response accuracy is still a relevant metric for older children. Older children also make mistakes in inhibition tasks (Dowsett & Livesey, 2000), and a speed-accuracy tradeoff can be observed for elementary school children (Best et al., 2011), that is, lower response accuracies due to hasty responses (Heitz, 2014). Even though constant improvement of inhibitory control during childhood is a stable pattern, large interindividual differences remain in each age group (Williams et al., 1999). Especially around the age of five, when many children reach a high accuracy in inhibitory control tasks, weaker inhibitory control may still translate into lower response accuracy. Conversely, in 4-year-olds, many of whom still make mistakes, the additional metric of response latency might be needed to differentiate between high-performers. These considerations emphasize the need for a test that measures efficient inhibitory control at the preschool age. Response latency and response accuracy need to be considered for children between the ages of 3 and 6 to obtain a thorough estimation of their inhibitory control ability.

Specific tests have been constructed for children because tests assessing inhibitory control in adults are often not suited for children (Berger et al., 2000). The Day-Night test (Gerstadt et al., 1994; for a review, see Montgomery & Koeltzow, 2010) and the Silly Sound Stroop task (Willoughby et al., 2011) are typically used to assess inhibitory control in 3- to 6-year-olds. In the Day-Night test, children are presented with a white card showing a bright sun or a dark card showing a moon and stars. Their task is to say *night* when they have been presented with the sun-card and *day* when they have been presented with the night-card. In each scenario, the prepotent congruent response (sun-card: *day*, night-card: *night*) has to be suppressed in favor of its incongruent opposite. Similarly, in the Silly Sound Stroop task, children are instructed to incongruently meow when they are presented with the picture of a dog and to bark when they see a cat. Whereas these tests are more engaging for younger children, their non-digital implementations lack a way to accurately measure response latency. Conversely, tests that are implemented digitally sometimes lack a playful element. In a study by Ikeda et al. (2014), for example, 3- to 12-year-olds were asked to categorize presented big circles as small and small circles as big in the inhibition condition. While their task was conducted digitally, the experimenter had to monitor the children's attentional focus.

The presented tests are only examples of typical inhibitory control tests (for more extensive lists of measures, see Carlson, 2005; Garon et al., 2008). Nevertheless, they highlight the challenge of creating tests for young children. Recent research moved towards digitalizing playful tasks with tablets (Willoughby & Blair, 2016). Despite computerized implementations, however, some studies continued to only focus on response accuracy (e.g., Verhagen et al., 2017) even though the integration of response latency became possible. Researchers have only recently begun to use both response accuracy and response latency when assessing inhibitory control in young children. Some studies have already demonstrated the benefits of this approach. Magnus et al. (2019) showed that integrating response latency into an item-level factor analysis model in addition to response accuracy yielded more precise measurements of inhibitory control compared to response accuracy alone. Magnus et al. used a bifactor-like model that integrated response accuracy and latency, which both loaded on the ability factor that reflected inhibitory control. In addition, the model contained an orthogonal factor that accounted for variance in latencies that was unrelated to inhibitory control. Specifically, the authors found that for two out of three computerized inhibitory control tasks, including response latency into the analyses resulted in a more accurate measurement of inhibitory control ability. Although no significant benefit of this approach was found for the Spatial Conflict Arrow task, the computerized Silly Sound Stroop task and the Animal Go/No-Go task benefitted from incorporating accuracy and latency, as indicated by lower standard errors of the factor scores. Moreover, Camerota et al. (2019) found that using response latency and response accuracy as measures of inhibitory control improved the prediction of academic and behavioral outcomes such as math skills or social skills from inhibitory control. Apart from these studies, however, no thorough psychometric evaluation of a computerized inhibitory control task for young children has been conducted that concurrently uses response latency and response accuracy.

A test that is engaging for children and enables both measurements is the computerized Pointing-Stroop task (cPST; Berger et al., 2000). The cPST is a digitally implemented Stroop task that produces incongruence via a visual-auditory incompatibility of the presented stimuli. Similar to the manual Silly Sound Stroop task introduced by Willoughby et al. (2011), children are presented with an image of a dog and an image of a cat on a touchscreen in every trial. Their task is to tap on the cat when they hear a *meow* and to tap on the dog when they hear a *woof* in the congruent block. In the incongruent block, they are asked to tap on the dog when they hear a *meow* and to tap on the cat when they hear the *woof*. In the original cPST, children were first presented with 16 congruent items before answering 16 incongruent items.

Berger et al. (2000) evaluated the cPST with 33 five-year-olds and found that the children needed more time responding to the incongruent trials compared to the congruent trials, indicating that the cPST is a valid test of inhibitory control in terms of response latency. However, response accuracy did not differentiate between children because of a ceiling effect. This ceiling effect, however, was to be expected, given that 5-year-olds have been found to respond highly accurately in inhibitory control tasks aimed at young children (Roebers, 2017). Despite these findings with the cPST, which suggest that response latency can be assessed reliably even in younger children with appropriate digitalized tasks, inhibitory control has continued to be assessed only in terms of response accuracy (e.g., Obradović, 2010). Although response accuracy is an important metric of response inhibition especially for young children (Paap, 2019), response latency might add valuable diagnostic information to the assessment of inhibitory control, especially when the test targets a broader age range from 3 to 6.

**Rationale of the Present Study**

The present study used a shortened version of the cPST (Berger et al., 2000) to psychometrically evaluate the diagnostic value of using both response accuracy and response latency as a measure of inhibitory control in young children. Previous research on inhibitory control in young children has primarily relied on response accuracy in Stroop-like tasks (e.g. Rhoades et al., 2009; Verhagen et al., 2017; Willoughby et al., 2012). Berger et al. (2000) used the cPST to show that children's response latency can also be used for a valid assessment of inhibitory control in young children. Recent studies suggest that using response accuracy and response latency can yield a more accurate measure of inhibitory control (e.g., Magnus et al., 2019). We aimed to extend these findings by investigating the extent that response latency and accuracy each provide complementary diagnostic information about inhibitory control in young children and the extent that using response latency adds incremental value to using only accuracy for the assessment of inhibitory control. Furthermore, we sought to examine aspects of the validity of this diagnostic approach of assessing inhibitory control.

Assessing response latency and accuracy may improve the measurement of inhibitory control and compensate for shortcomings when using just one metric such as ceiling effects for response accuracy or neglecting a speed-accuracy tradeoff. Importantly, we aim to evaluate the task for a sample of children that spans the complete age range of kindergarten, that is, for 3- to 6-year-olds. Given the large leaps in the development of executive functions, including inhibitory control, in young children (Carlson & Moses, 2001; Kochanska et al., 1996), evaluating a test across a broader age span is critical because capabilities in inhibitory control might express themselves differently for older compared to younger children (Petersen et al., 2016; Roebers, 2017).

We created an eight-item version of the cPST to allow for a more time-economical measurement that will not overtax young children's ability to maintain an on-task focus. This is especially important when assessing inhibitory control in the context of a larger test battery, as it's often done. We also estimated explanatory item response models (De Boeck & Wilson, 2004) to examine the construct validity of the cPST. We used (Generalized) Linear Mixed Models (GLMM/LMM) for the explanatory item response analyses. This trial-wise analysis approach allows taking both person- and item-level effects into consideration in one model and is suitable for handling the clustered (crossed) data structure (items nested within participants and participants nested within items, Baayen et al., 2008).

For these analyses, we expected that the item difficulties (response accuracy) and time intensities (response latency) of the cPST would vary systematically between congruent and incongruent trials. Specifically, we expected children to make more mistakes and take longer to respond to incongruent compared to congruent items (Hypothesis 1a and 1b). Contrary to past research, the approach of estimating explanatory item response models allows for an item-level analysis, which opens the possibility of accounting for sequence effects in inhibition tasks. Previous research has shown a practice effect in response inhibition tasks (MacLeod & Dunbar, 1988; Zhao et al., 2018), that is, a positive effect of the position of each item in the test (the position at which the test item occurs in the test block) on response accuracy and a negative effect of item position on response latency. We expected a practice effect, especially in incongruent trials, for both response accuracy (Hypothesis 2a) and response latency (Hypothesis 2b). Thus, the basic idea was that if the items of the cPST assess individual differences in inhibitory control, the item features *congruency* and *item position* should have an effect on empirical item difficulties (for response accuracy) and time intensity (for response latency).

We relied on the typical three-step procedure to evaluate the hypothesized explanatory item response model (Wilson & DeBoeck, 2004; Hartig et al., 2012; Schindler et al., 2018) and thus the construct validity of the cPST. The method enables the estimation of specific effects that item characteristics (in this case congruency and position) have on the item difficulty. It allows calculating an expected item difficulty (estimated difficulty) for each item. A substantial positive correlation between the empirical (observed) and the model-based estimated item difficulties for accuracy (Hypothesis 3a) and time intensities for latency (Hypothesis 3b) would establish that congruency and item position affect item difficulty in terms of both accuracy and response latency. This would corroborate the construct validity of the short version of the cPST with response latency and accuracy as measures of inhibitory control for kindergarten children.

Given that children in the older ages of kindergarten tend to make almost no mistakes in inhibitory control tests that were adapted for young children (Berger et al., 2000; Roebers, 2017), we expected that the metric of response accuracy is more suitable for assessing inhibitory control in children younger than 5, whereas response latency should be a valid measure across all age groups. This assumption would be supported by an interaction between the participants' age and the effect of congruency as predictors for the children's response accuracy (Hypothesis 4). A significant interaction would indicate that the effect of item incongruence on response accuracy differs depending on participants' age. No such interaction was expected for response latency.

Additionally, we examined the cPST's construct validity by examining the relationship between the efficiency of inhibitory control measured with the cPST and fluid intelligence. We expected children who perform efficiently in the cPST to reach higher intelligence scores because of the theoretical (Dempster, 1991) and empirical (e.g., Kane & Engle, 2002) links between intelligence and inhibitory control. Given that performance in incongruent items is commonly used to measure inhibitory control (e.g., Willoughby et al., 2012), we followed the same procedure. We expected a nonadditive effect of accuracy and latency in the incongruent items of the cPST on fluid intelligence. Children who respond accurately and quickly should show a particularly high fluid intelligence (Hypothesis 5). This pattern of effects would further corroborate the validity of using both response latency and accuracy as indicators for inhibitory control. Children capable of efficient inhibition would then display higher capability in a metric closely related to inhibitory control, namely fluid intelligence.

## Method

### Participants

Participants were 135 children (71 girls) from Switzerland and Germany recruited via childcare facilities, the experiment's website, and local newspapers. Their mean age was 4.24 years (*SD* = 7.42 months) ranging from 2.92 to 6.50 years. Sociodemographic data were collected via a parent's questionnaire. The sample included 42 bilingual children. Of the 135 children, 33 spoke High German, 31 spoke Bernese German, 22 spoke Basel German, 22 spoke French, 10 spoke Italian and 17 spoke Turkish as their dominant language. The average socioeconomic status, measured by the highest education of any parent, was high. Eighty-eight children (out of 115 valid data points; 77%) had at least one parent who held a university degree. For 20 children, parents failed to complete the parental survey.

### Procedure

Data were collected in the context of a cross-national project aiming to examine the cognitive, metacognitive, linguistic, and socio-emotional development of mono- and bilingual 3- to 6-year-olds (https://www.crocodile-study.ch). Testing was conducted between November, 2019 and April, 2020. The children were tested in their childcare facilities, at home, or in dedicated lab spaces at universities. The sessions were conducted in individual settings with the research assistant, the child, and the parent or an educator being present if necessary. Testing took part at two different days for each child with 40 to 60 min sessions per day. Most of the testing, including the cPST, was conducted on a 14" Windows convertible. The children were led through the experiment by a crocodile named Sammy who invited them on a virtual treasure hunt. Including story, instruction, and practice, the cPST took 5 min on average. Story, instructions, and practice lasted 2.5 minutes.

## Measures

### *Computerized Pointing-Stroop Task*

The shortened version of the cPST was part of the test battery used in the project and was presented at different time points during the study, the order of which was counterbalanced across participants. As part of the story, Sammy and the child met a dolphin that was occupied with learning what dogs and cats sound like. The task consisted of two blocks. In both blocks, children were presented with an image of a cat and an image of a dog next to each other (position randomized between trials), and they simultaneously heard a bark or a meow. In the first block (congruent block), children were instructed to press the image matching the animal sound, that is, the cat when hearing a meow and the dog when hearing a bark. In the second block (incongruent block), children were instructed to press the image contradicting the animal sound, that is, pressing the dog when hearing a meow and pressing the cat when hearing a bark. Each block consisted of four trials. Before each block, two practice trials needed to be solved correctly by the child. When an incorrect answer was given, the practice trial was repeated. The story and instructions were presented in the language children were most proficient in.

### *Fluid Intelligence*

The Categories subscale of the Snijders-Oomen Non-Verbal Intelligence Test (SON-R 2 ½ – 7; Tellegen et al., 2006) was used as an indicator of the children's non-verbal fluid intelligence. It aims to assess abstract thinking capability in young children and was presented offline (i.e., non-computerized) as part of the project's test battery at different time points for each child, depending on the randomization. In a first block, children were instructed to sort four to six cards depicting objects into predefined categories. In a second block, children were instructed to select one card from a pool of five cards that shared a common feature with three other predefined cards.

**Analytic Strategy**

The analytic strategy encompassed the estimation of explanatory item response models for Hypotheses 1-3, (generalized) linear mixed models for Hypothesis 4 and a linear model for Hypothesis 5. Data and R-code for the reported analyses are available at the repository of the Open Science Framework (https://osf.io/nckpb/?view_only=f23b504b4db5445488222aa65fe98bec).

**Explanatory Item Response Models**. In Step 1, item difficulties were estimated using a one-parameter logistic item response model (1PL model) for the response accuracy (Hartig et al., 2012), implemented as a Generalized Linear Mixed Model (GLMM; Dixon, 2008) with a logit-link. Likewise, time intensities were estimated using an analogous Linear Mixed Model (LMM; Baayen et al., 2008) for the (log-)transformed response latency (Van Breukelen, 2005). The items were dummy-coded with Item 1 from the congruent block as reference category and included as predictors (fixed effects) for response accuracy or log-transformed response latency as dependent variables. The effects of these variables represent the item parameter of the 1PL model and the analogous LMM model, that is, differences in item difficulty (response accuracy) and time intensity (response latency). The intercept could vary randomly between participants (random intercept), representing the person parameter. Formula 1 displays the model equation for the LMM estimated in Step 1 for response latency.

(1) $\ln(latency_{ij}) = b_{0j} + b_{1j} X_{1ij} + b_{2j} X_{2ij} + ... + b_{7j} X_{7ij} + r_{ij}$.

$b_{0j} = g_{00} + u_{0j}$ *Random coefficient*, $u_{0j}$: person parameter

$b_{1j} = g_{10}$

... *Fixed coefficients*: item parameters

$b_{7j} = g_{(k-1)0}$

The GLMM for response accuracy as outcome variable was set up in the same way as a random-intercept model with fixed item effects.

In Step 2, item difficulties were modeled in a Logistic Linear Test Model (LLTM; Fischer, 1974) and time intensities in an analogous linear mixed model (Van Breukelen, 2005) as a function of congruency (0 = congruent, 1 = incongruent) and presentation position in the block (grand-mean centered). Again, a GLMM with a logit-link was estimated for response accuracy and an LMM for the (log-)transformed response latencies. Both models included a random intercept and fixed effects of congruency and item position. Formula 2 displays the model equation for the LMM for latencies. The GLMM for response accuracy was set up in the same way.

(2) $\ln(latency_{ij}) = b_{0j} + b_{1j}\ (congruency)_{ij} + b_{2j}\ (item\ position)_{ij} + r_{ij}$.

$b_{0j} = g_{00} + u_{0j}$ Random coefficient, $u_{0j}$: person parameter

$b_{1j} = g_{10}$ Fixed effect of congruency

$b_{2j} = g_{20}$ Fixed effect of item position

The intercepts and slopes from the Step 2 models were then used to derive estimated item difficulties and time intensities for each item. In Step 3, estimated item difficulties and time intensities obtained in Step 2 were correlated with the empirical item difficulties and time intensities obtained in Step 1 to determine how well congruency and item position predict the item difficulties and time intensities.

**Age Effect.** To examine the effect of the participants' age on the congruency effect, we extended the GLMM for response accuracy and the LMM for the (log-)transformed response latency estimated in Step 2 by adding age as a grand-mean centered predictor and its interaction with congruency.

**Relationship with Intelligence**. To examine the relationship between participants' inhibitory control and their SON-R scores, we analyzed the incongruent trials of the cPST, given that incongruence-generating trials are commonly used as indication for inhibitory capabilities (e.g., Willoughby et al., 2012). We used the children's raw SON-R scores because of the age sensitivity of measurements in inhibitory control (Roebers, 2017). We estimated a moderated linear regression model (Aiken & West, 1991) with mean response accuracy (z-standardized), mean log-transformed response latency (z-standardized), and their interaction as predictor variables and the children's SON-R scores as dependent variable.

## Results

The data of three children were excluded from the analyses because they received help from the research assistant, parent, or educator, objectively were not able to comprehend the instructions (responded with two fingers simultaneously), or refused to carry out the task. Children with an error rate of over .70 in the congruent condition were assumed to have failed to learn the task as the low accuracy in the congruent block could not be explained by a low capability in inhibitory control (see Davidson et al., 2006, p. 2041). They were also excluded ($n$ = 7), resulting in a sample of 125 participants. Response latencies from incorrect responses were included in all response latency analyses. The LMMs/GLMMs were estimated with the packages lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages for the statistical software R (R Core Team, 2019; Version 4.0.3). The moderated regression analysis for analyzing the relationships of inhibitory control with intelligence and the moderating effect of age on the effect of congruency was conducted with the package interactions (Long, 2019). Parameters were estimated with Restricted Maximum Likelihood (REML). All significance tests were based on a Type I error probability of .05.

## Explanatory Item Response Models

We will report the empirical item difficulties (Step 1) and descriptive item statistics, the results from the LLTM and the LLT-analogous model (Step 2), and correlations between empirical and estimated item difficulties (Step 3) are reported separately for response accuracy and latency.

### *Response Accuracy*

**Step 1.** The 1PL model for logit-transformed response accuracy as dependent variable revealed that the items of the cPST were easy to solve for the children. Empirical item difficulties (log-odds for solving a specific item across participants) are depicted in the left panel of the person-item map in Figure 1 (a value of 4 equals 98% correct responses). Person abilities (log-odds for providing a correct response across items) are depicted in the right panel of the person-item map in Figure 1. In comparison to congruent items, incongruent items seem to be suitable to differentiate between children in the lower-ability range.

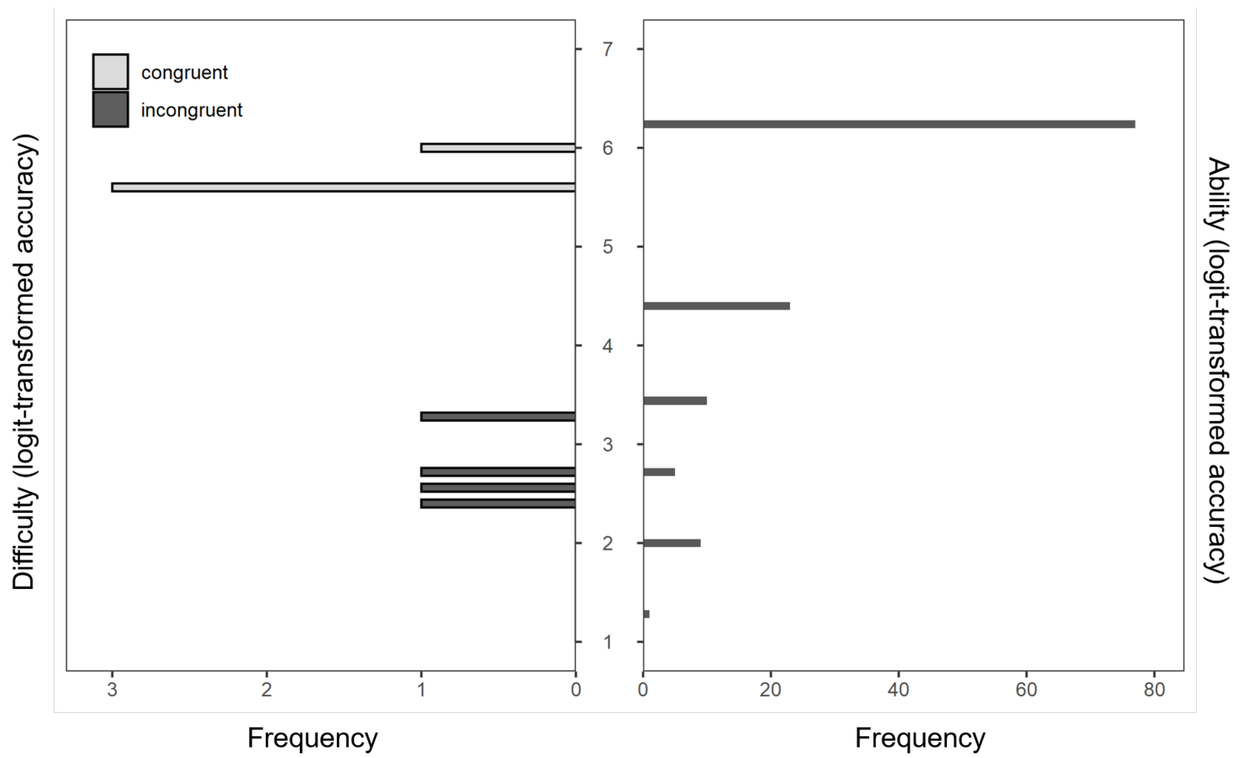
**Step 2.** The parameter estimates for the fixed and random effects of the LLTM with logit-transformed response accuracy as dependent variable are provided in Table 1 (Response Accuracy). The main effect for incongruence reached significance. In line with Hypothesis 1a, children's response accuracy was lower in incongruent trials ($P = .94$, $SE = .02$) compared to congruent trials ($P = 1.00$, $SE = .00$; $\beta = -2.97$; $z = -7.52$, $p < .001$). The main effect for item position was not significant ($\beta = -0.22$; $z = -1.82$, $p = .069$). Thus, Hypothesis 2a did not receive support, possibly because children responded quite accurately from the start (ceiling effect).

**Step 3.** The estimated item difficulties correlated strongly and positively, $r(6) = .98$, $p < .001$, with the empirical item difficulties. The estimated item difficulties explained 97%

of the variance in the empirical item difficulties (Figure 2), indicating an excellent fit of the explanatory item response model. Thus, Hypothesis 3a was supported.

**Figure 1**

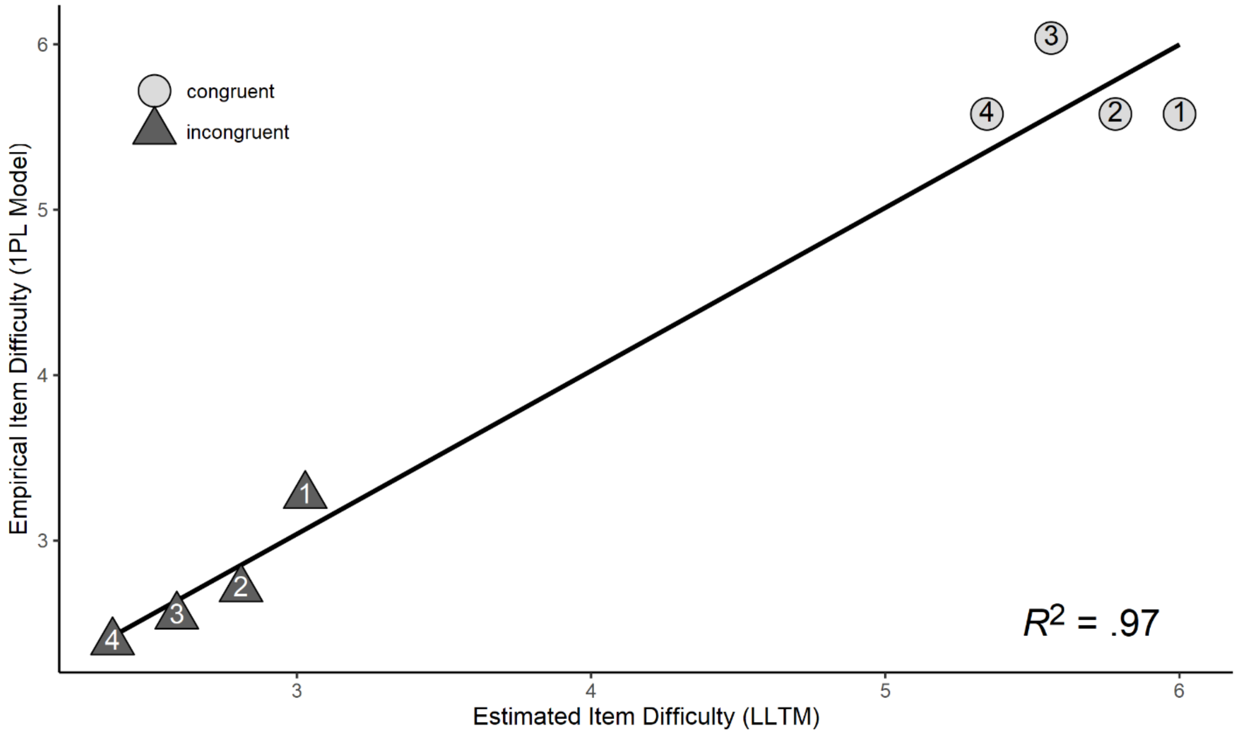
*Person-Item Map for Logit-transformed Response Accuracy*




*Note.* The left panel displays the distribution of item difficulties, that is, the logit-transformed response accuracy. The right panel displays the person abilities.

**Figure 2**

*Variance in Empirically Observed Item Difficulties Explained by Item Difficulties Predicted from Item Characteristics for Response Accuracy in the Computerized Pointing-Stroop Task*




*Note*. Digits in the data points indicate the item position within the respective block. The y-axis displays the empirical difficulty, that is, the log-odds for solving a specific item across participants, and the x-axis displays the estimated difficulty based on the item features congruency and item position.

**Table 1**

*Fixed Effects and Variance Components in the LLTM for Response Accuracy and LLT-Analogous Models for Response Latency in the Computerized Pointing-Stroop Task.*

| | Response Accuracy | Response Latency |
|---|---|---|
| Parameter | β (*SE*) | β (*SE*) |
| Fixed Effects | | |
| Intercept | 5.673 (0.56)*** | 0.654 (0.03)*** |
| Congruency | -2.972 (0.40)*** | 0.162 (0.02)*** |
| Position | -0.218 (0.12) | -0.116 (0.01)*** |
| Variance Components | | |
| Participant | 4.566 (2.14) | 0.090 (0.30) |

*Note*. Congruency (0 = congruent, 1 = incongruent). Position (grand-mean centered).

* $p < .05$. ** $p < 0.05$. *** $p < 0.01$. (two tailed)

***Response Latency***

**Step 1.** The linear model for log-transformed response latency as dependent variable revealed substantial overlap of the distributions of empirical time intensities (log-transformed latencies for providing a response to a specific item across participants) and person abilities (log-transformed latencies for providing a response across items). This finding suggests that the set of

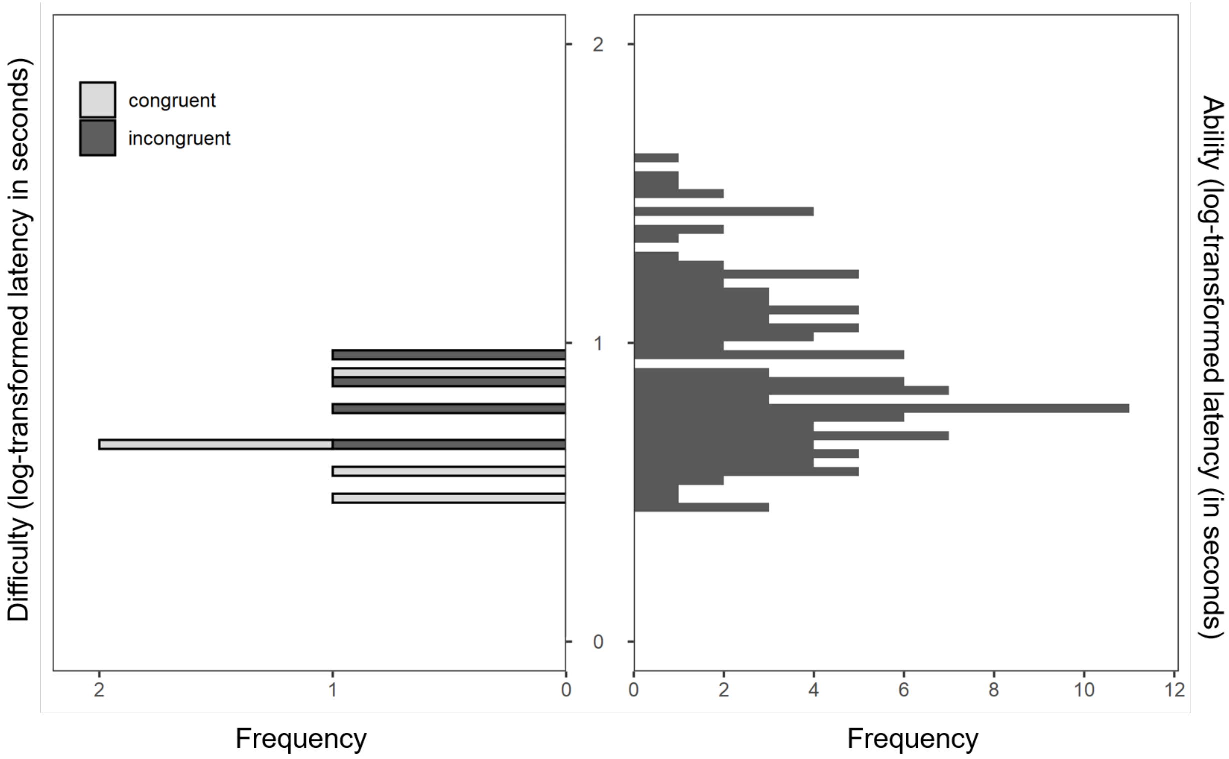
items measures individual differences across most of the range of person abilities, except for low performing children. The distribution of item difficulties is depicted in the left panel of the person-item map in Figure 3 and the distribution of person abilities in the right panel (a log-transformed response latency of 1 equals 2.72 seconds).

**Step 2.** The parameter estimates for the fixed and random effects of the LLT-analogous model with log-transformed response latency as dependent variable are provided in Table 1 (Response Latency). The main effect for incongruence reached significance. In line with Hypothesis 1b, children needed more time for incongruent items (*M* = 2.26 s, *SE* = 0.07) compared to congruent items (*M* = 1.92 s, *SE* = 0.06; β = 0.16; $t(873) = 6.95$, $p < .001$). The main effect for position also reached significance. In line with Hypothesis 2b, children responded faster to items occurring later in the block (β = -0.12; $t(873) = -11.12$, $p < .001$), indicating a practice effect. To investigate the relationship of response accuracy and response latency, we estimated two additional models: In the first model, response accuracy of each trial was included as fixed and random effect over participants. For the second model, the accuracy of the previous trial was included as fixed and random effect over participants. Neither predictor had a significant influence on response latency. Results and syntax from the exploratory analyses are available in the repository of the Open Science Framework (https://osf.io/nckpb/?view_only=f23b504b4db5445488222aa65fe98bec).

**Step 3.** The estimated time intensities derived from the Step 2 model correlated strongly and positively, $r(6) = .97$, $p < .001$, with the empirical time intensities. The estimated time intensities explained 93% of the variance in the empirical time intensities (Figure 4), indicating again an excellent fit of the explanatory item response model to the observed time intensities. Thus, Hypothesis 3b was supported.

**Figure 3**

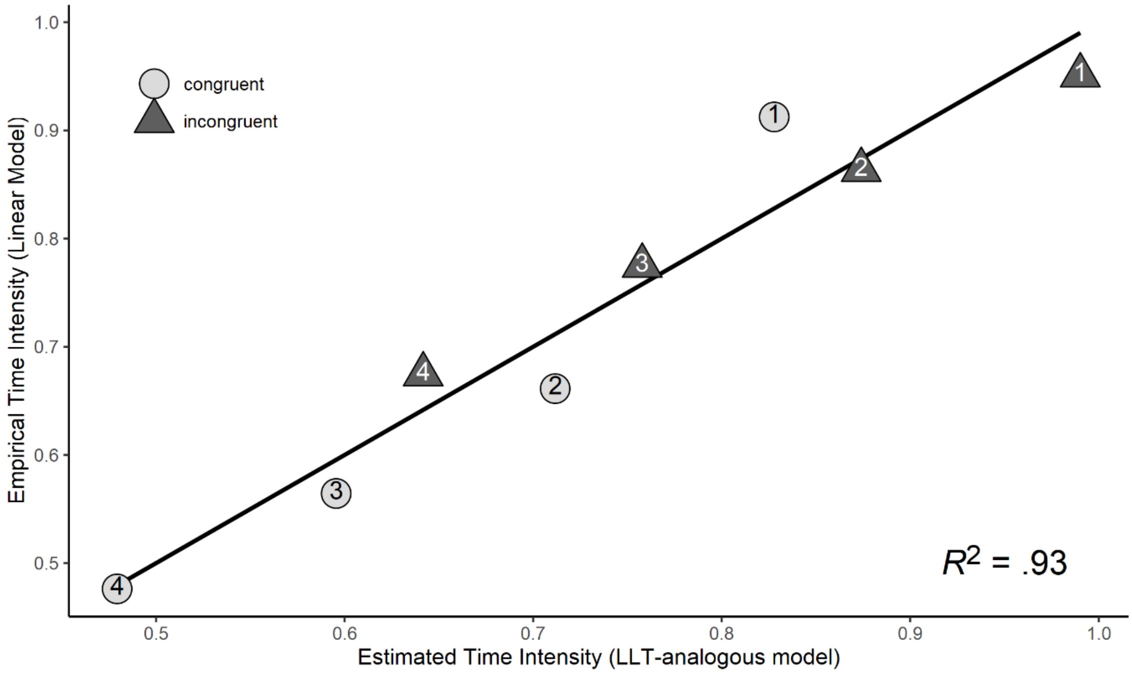
*Person-Item Map for Log-transformed Response Latencies*




*Note.* The left panel displays the distribution of item time intensities, that is, the log-transformed response latencies. The right panel displays the person abilities.

**Figure 4**

*Variance in Empirically Observed Item Time Intensities Explained by Time Intensities Predicted from Item Characteristics for Response Latency in the Computerized Pointing-Stroop Task*



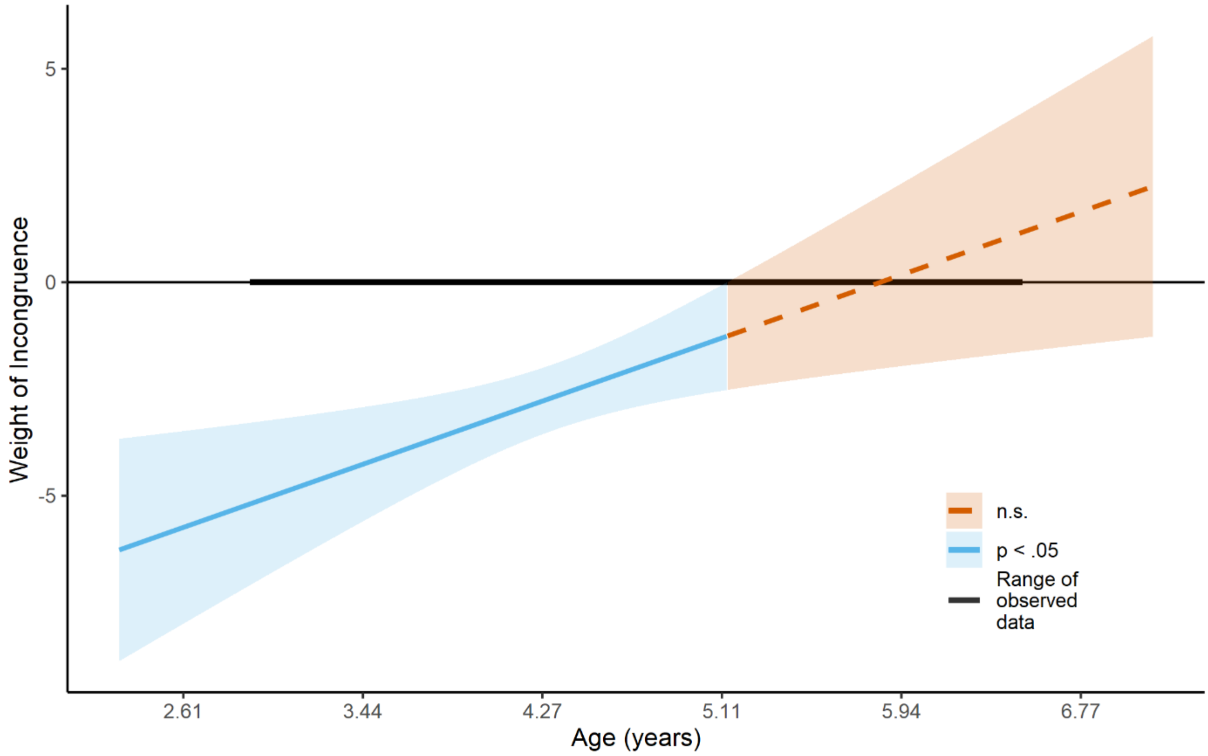

*Note*. Digits in the data points indicate the item position within the respective block. The y-axis displays the empirical time intensity, that is, the log-transformed latencies for providing a response to a specific item across participants, and the x-axis displays the estimated time intensity based on the item features congruency and item position.

**Congruency and Age**

The interaction between participants' age and congruency reached significance for response accuracy (β = 1.08; $z = 2.82$, $p = .005$) but not for response latency (β = 0.04; $t(872) = 1.66$, $p = .098$). Simple slope analysis revealed that in terms of accuracy, incongruence had the strongest negative effect in young children but was no longer a significant predictor in children older than 5 years and 2 months (10.30 months above the mean; Figure 5). Thus, Hypothesis 4 was supported.

**Figure 5**

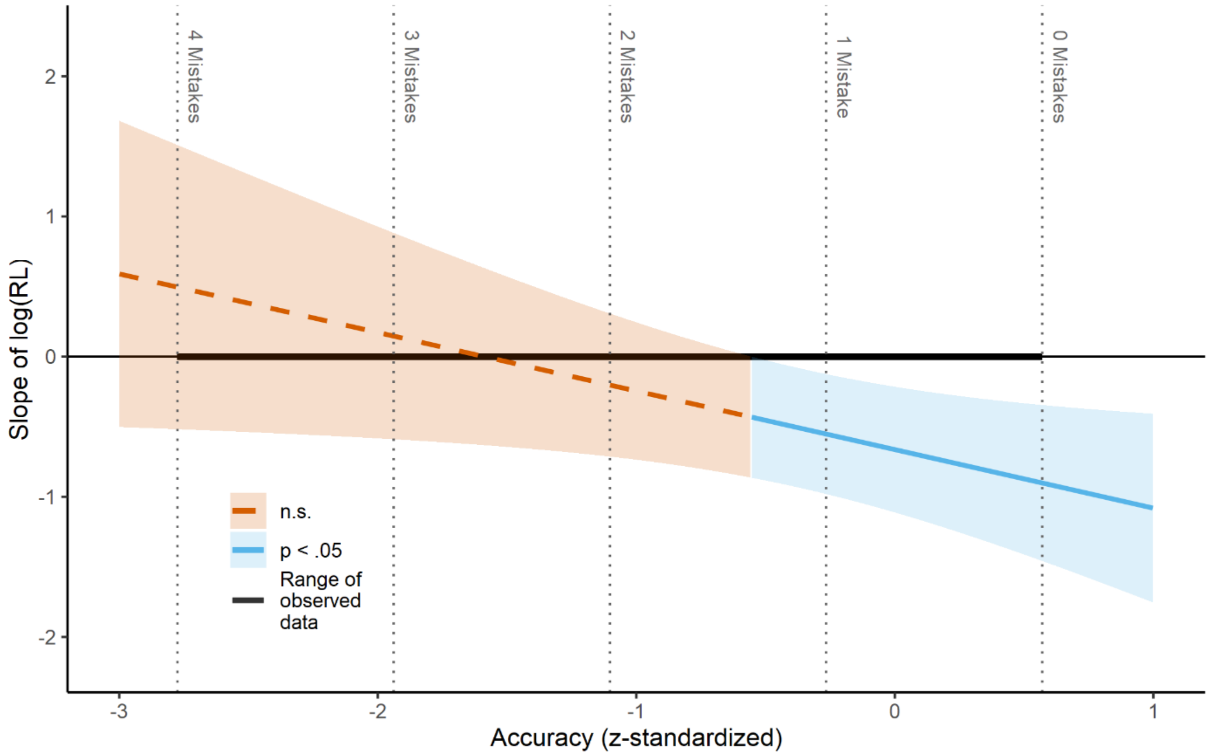
*Johnson-Neyman Plot: Effect of Incongruence on Response Accuracy Based on Participants' Age*




*Note.* The effect of incongruence on participants' accuracy in the cPST (y-axis) moderated by participants' age (x-axis). The colored area around the regression line represents the 95% confidence bands determining the region of significance.

**Figure 6**

*Johnson-Neyman Plot: Slope of (log) Response Latency Depending on Accuracy*




:*Note*. The effect of log-transformed response latency in incongruent items of the cPST on SON-R scores (y-axis) moderated by response accuracy (x-axis). The colored area around the regression line represents the 95% confidence bands determining the region of significance.

## Inhibitory Control and Intelligence

We also examined whether children who respond faster and more accurately would show a higher intelligence. In support of this assumption, the log-transformed response latencies in incongruent trials significantly predicted the children's SON-R scores (β = –0.66; $t(106) = -2.93$, $p = .004$). Children who responded faster during the incongruent trials of the cPST showed a higher SON-R score. Although the effect for accuracy suggested a positive relationship between the cPST and intelligence (β = .26; $t(106) = 1.22$, $p = .226$), the effect was not significant. However, the interaction of accuracy and response latency in the cPST reached significance (β = -.42; $t(106) = -2.17$, $p = .032$). In line with Hypothesis 5, children who responded quickly *and* with high accuracy scored highest in the SON-R (Figure 6). In addition to the significant main effect for response latency, accuracy significantly predicted the SON-R score in faster-responding children (Figure 6).

## Discussion

In this study, we examined the diagnostic utility of using both response accuracy and response latency to assess inhibitory control in a simple inhibition task, the cPST (Berger et al., 2000), in young children aged 3-6. Explanatory item response models for response accuracy and response latency indicated that the observed item difficulties could be almost perfectly explained by congruence vs. incongruence and item position. As expected, incongruence had a negative impact on the children's response accuracy, and it slowed down their responses. Congruency affected the participants performance in terms of accuracy until the age of 5 years and 2 months but not beyond that age. Participants who answered quickly and with high accuracy showed the highest fluid intelligence scores. The reported findings suggest that the items of the cPST validly assess individual differences in inhibitory control when response latency and accuracy are considered. The construct validity is supported by the findings that theoretically relevant item characteristics, that is, incongruence and item position, affected item difficulty and time intensities in a well-interpretable manner. The age sensitivity of the cPST and the finding that children who performed efficiently (i.e., quickly and accurately) displayed a higher fluid intelligence may be considered as further evidence for the construct validity.

The findings show that a simple inhibition task such as the cPST is a suitable instrument to measure inhibitory control in children across the whole age span of kindergarten and a broad ability range, provided that response accuracy and response latency are used. Using response accuracy may be important for differentiation between younger children and in the lower-ability range (Figure 1), whereas response latency may differentiate between older children and those in the higher ability range (Figure 3). A study by Camerota et al. (2020) that investigated 7-year-olds supports these conclusions. The authors integrated response latency and accuracy in computerized inhibitory control tasks by applying a bifactor model to item-level accuracy and response latency data in the Hearts and Flowers task. Results from their analysis suggest that especially for high performing children, response latency should be used in addition to response accuracy when estimating the inhibitory control performance. An arbitrary selection of either metric may lead to loss in validity of inhibitory control measures.

This is especially true for the age range of 3- to 6-years-olds because the more informative metric for inhibitory control might switch from response accuracy to response latency (Roebers, 2017). Although some researchers argue that response accuracy might be the more appropriate metric for assessing inhibitory control in young children (Paap, 2019), at a certain age, high performers reach a ceiling in child-oriented inhibitory control tasks (e.g., Berger et al., 2000). This ceiling effect was mirrored in the presented results. Incongruence no longer significantly predicted response accuracy in children older than 5 years and 2 months in the cPST, indicating that response accuracy loses its diagnostic value for assessing inhibitory control, which in turn heightens the importance of response latency as an inhibitory control measure for older children. Petersen et al. (2016) found that on average, inhibitory control tasks for children are only suitable for an age range of three years, depending on the target age**:** The Shape Stroop task (Kochanska et al., 1997), for example, yields valid measurements of inhibitory control in children as young as 1.5 years, but loses its validity after the age of 3.5. Other tests such as the Simon Says task (Strommen, 1973) validly measure inhibitory control until the age of 7, but should not be used before the age of 4.5 years. As the authors argue, this is primarily the case when only response accuracy is considered. The present study emphasizes that using both response accuracy as well as response latency is useful, especially when children over a broader age range are tested, such as in a longitudinal study.

These findings might help to clarify (and simplify) the question of which type of data are to be used in inhibitory control measures. Sometimes this decision is justified by the participants' age. For example, Diamond et al. (2007) argued that for their sample of 5-years-olds, accuracy is the more reliable metric. In contrast, Ursache and Raver (2015) argued that the high accuracy in their sample of 9- to 12-years-olds allows for no analysis of response accuracy. However, combining response accuracy and response latency to measure inhibitory control provides an increment in diagnostic information for younger children, as demonstrated in the present study, and for older children (Camerota et al., 2019; Camerota et al., 2020). Therefore, using response accuracy and response latency in the assessment of inhibitory control seems to be advisable in children of any age, starting from the age of 3.

The shortened version of the cPST that was used in the present study has proven its utility for assessing efficient inhibitory control, that is, a quick and accurate response inhibition. Its non-verbal implementation and easy and engaging procedure make the test especially suitable for young children and situations in which non-verbal tasks are crucial, for example, in research with bilingual children. Moreover, this shortened adaptation of the cPST is a more economic version that may keep the children more engaged when advancing to the incongruent block because of its brevity. Children may experience less cognitive fatigue from four items in the congruent block (compared to 16 items in the original version). This makes the short version of the cPST particularly appropriate for inclusion in longer test batteries or extended test sessions with multiple tasks.

Its informative and clear results notwithstanding, this study also has limitations. One potential limitation is due to the nonrandom sample employed in the present study. The sample was clearly selective because of the restricted variance in the socioeconomic status of the children's families. Whereas this hurts the generalizability of the results, the implications for the cPST are likely to be minor: Studies have found a negative relationship between socioeconomic status and inhibitory control (Kałamała et al., 2020; Lipina et al., 2013). One might expect an overall lower accuracy and, consequently, a larger interindividual variance of the accuracy in the cPST in a sample that includes more children coming from families with a lower socioeconomic status. A more diverse sample in terms of SES would likely result in a higher discriminatory value of response accuracy in the cPST. Furthermore, the diverse sample of children from different countries and varying cultural backgrounds and an age range spanning the complete kindergarten age contributed to the ecological validity of the study. A second potential limitation is that our method only examined inhibitory control without assessing the other two executive abilities, working memory and cognitive flexibility. Intelligence is a core correlate of inhibitory control, as established in our study, but the exact nature of the relationship between these constructs has been a point of contention. Although a relationship has been established in several studies (Duan et al., 2010; Friedman et al., 2006; Yücel et al., 2012), inhibitory control shares considerable variance with other executive functions when predicting intelligence (Duan et al., 2010; Friedman et al., 2006). Future studies should address the relationships of efficient inhibitory control with intelligence while controlling for other executive functions. As a third limitation, a full investigation of potential benefits of measuring inhibitory control with response accuracy and latency would also require examining the relationship of the construct with real-world outcomes of inhibitory control (e.g., Camerota et al., 2019), which is certainly a question that would be worth pursuing in future research. Finally, the current study demonstrates that using both response accuracy and latency yields incremental value opposed to using only one. The next step would consist of integrating both metrics in a way that produces scores that are representative of participants' inhibitory control. Common methods of integrating response accuracy and response latency frequently work with aggregated data from incongruent trials (Liesefeld & Janczyk, 2019; Vandierendonck, 2017). These methods however, do not allow controlling for important factors such as the base response speed of participants.

To conclude, complementing the findings from factor-analytic studies (Camerota et al., 2019; Magnus et al., 2019), the results of the present study demonstrate the feasibility and validity of using response accuracy and response latency in a simple inhibition task to assess inhibitory control in children at the age of 3 to 6. Moreover, the results suggest that when using a task such as the cPST (Berger et al., 2000), the approach requires only a relatively small number of items. These results are of practical value for researchers who wish to assess inhibitory control in kindergarten children in a valid and economical way.

## References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Berger, A., Jones, L., Rothbart, M. K., & Posner, M. I. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. https://doi.org/10.3758/BF03207798

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21*(4), 327–336. https://doi.org/10.1016/j.lindif.2011.01.007

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin, 143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*(1), 48-71. https://doi.org/10.1016/0273-2297(90)90004-N

Camerota, M., Willoughby, M. T., & Blair, C. B. (2019). Speed and accuracy on the Hearts and Flowers task interact to predict child outcomes. *Psychological Assessment, 31*(8), 995–1005. https://doi.org/10.1037/pas0000725

Camerota, M., Willoughby, M. T., Magnus, B. E., & Blair, C. B. (2020). Leveraging item accuracy and reaction time to improve measurement of child executive function ability. *Psychological Assessment, 32*(12), 1118–1132. https://doi.org/10.1037/pas0000953

Campbell, K. L., Lustig, C., & Hasher, L. (Eds.). (2020). Aging and inhibition: The view ahead [Special issue]. *Psychology and Aging, 35*(5).

Carlson S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

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

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037-2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Boeck, P., & Wilson, M. (Eds.). (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. Springer. https://doi.org/10.1007/978-1-4757-3990-9

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157-173. https://doi.org/10.1016/0160-2896(91)90028-C

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development, 85*(5), 1777-1794. https://doi.org/10.1111/cdev.12237

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169539.003.0006

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*(5855), 1387–1388. https://doi.org/10.1126/science.1151148

Dixon, P. (2008). Models of accuracy in repeated measures designs. *Journal of Memory and Language, 59*(4), 447–456. https://doi.org/10.1016/j.jml.2007.11.004

Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology, 36*(2), 161–174. https://doi.org/10.1002/(SICI)1098-2302(200003)36:2<161::

Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year-old children. *Psychological Test and Assessment Modeling, 52*(4), 419–431.

Escobar, G. P., Kalashnikova, M., & Escudero, P. (2018). Vocabulary matters! The relationship between verbal fluency and measures of inhibitory control in monolingual and bilingual children. *Journal of Experimental Child Psychology, 170*, 177–189.

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6*(6), 409–434. https://doi.org/10.1080/02699939208409696

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests* [Introduction to the theory of psychological tests]. Huber.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Garon, N. (2016). A review of hot executive functions in preschoolers. *Journal of Self-Regulation and Regulation*, 2, 57– 80. https://doi.org/10.11588/josar.2016.2.34354

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Hartig, J., Frey, A., Nold, G., & Klieme, E. (2012). An application of explanatory item response modeling for model-based proficiency scaling. *Educational and Psychological Measurement, 72*(4), 665–686. https://doi.org/10.1177/0013164411430707

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*: 150. https://doi.org/10.3389/fnins.2014.00150

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*(2), 107–129. https://doi.org/10.1016/0001-6918(67)90011-X

Ikeda, Y., Okuzumi, H., & Kokubun, M. (2014). Age-related trends of inhibitory control in Stroop-like big-small task in 3 to 12-year-old children and young adults. *Frontiers in Psychology, 5*: 227. https://doi.org/10.3389/fpsyg.2014.00227

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498–504. https://doi.org/10.1111/1467-7687.00307

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition, 204*, Article 104373. https://doi.org/10.1016/j.cognition.2020.104373

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*(4), 637-671. https://doi.org/10.3758/BF03196323

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263–277. https://doi.org/10.2307/1131849

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490-507. https://doi.org/10.2307/1131828

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Lipina, S., Segretin, S., Hermida, J., Prats, L., Fracchia, C., Camelo, J. L., & Colombo, J. (2013). Linking childhood poverty and cognition: Environmental mediators of non-verbal executive control in an Argentine sample. *Developmental Science, 16*(5), 697–

707. https://doi.org/10.1111/desc.12080

Long, J. A. (2019). *interactions: Comprehensive, user-friendly toolkit for probing interactions* (R package version 1.1.0) [Computer Software].

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 126–135. https://doi.org/10.1037/0278-7393.14.1.126

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*(7), 1296–1306. https://doi.org/10.1177/1073191117740953

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*(3), 308–330. https://doi.org/10.1016/j.dr.2010.07.001

Moriguchi, Y., & Hiraki, K. (2013). Prefrontal cortex and executive function in young children: A review of NIRS studies. *Frontiers in Human Neuroscience, 7*, 867. https://doi.org/10.3389/fnhum.2013.00867

Ng, F. F.-Y., Tamis-LeMonda, C. S., Yoshikawa, H., & Sze, I. N.-L. (2014). Inhibitory control in preschool predicts early math skills in first grade: Evidence from an ethnically diverse sample. *International Journal of Behavioral Development, 39*(2), 139–149. https://doi.org/10.1177/0165025414538558

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Müller, U., & Kerns, K. (2015). The development of executive function. In L. S. Liben, U. Müller, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (pp. 571–623). John Wiley & Sons. https://doi.org/10.1002/9781118963418.childpsy214

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences, 15*(10), 453-459.

Obradović, J. (2010). Effortful control and adaptive functioning of homeless children: Variable-focused and person-focused analyses [Editorial]. *Journal of Applied Developmental Psychology, 31*(2), 109–117. https://doi.org/10.1016/j.appdev.2009.09.004

Paap, K. R. (2019). The bilingual advantage debate: Quantity and quality of the evidence. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch34

Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., & Staples, A. D. (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review, 40*, 25–71. https://doi.org/10.1016/j.dr.2016.02.001

R Core Team (2019). *R: A language and environment for statistical computing* (Version 4.0.3) [Computer Software]. Retrieved from: https://www.R-project.org/

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social–emotional competence. *Journal of Applied Developmental Psychology, 30*(3), 310–320. https://doi.org/10.1016/j.appdev.2008.12.012

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain: A Journal of Neurology, 133*(1), 234–247. https://doi.org/10.1093/brain/awp269

Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology, 12*(1), 105-126. https://doi.org/10.1080/87565649609540642

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Simpson, A., & Carroll, D. J. (2019). Understanding early inhibitory development: Distinguishing two ways that children use inhibitory control. *Child Development, 90*(5), 1459-1473. https://doi.org/10.1111/cdev.13283

Strommen, E. A. (1973). Verbal self-regulation in a children's game: Impulsive errors on "Simon says.". *Child Development, 44*(4), 849–853. https://doi.org/10.2307/1127737

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(10), 1231–1238. https://doi.org/10.1097/00004583-200210000-00013

Tellegen, P. J., Laros, J. A., & Petermann, F. (2006). *Deutsche Normierung und Validierung des SON-R 2 ½ - 7* [Norming and validation of the SON-R 2 1/2 – 7 in German]. Hogrefe.

Ursache, A., & Raver, C. C. (2014). Trait and state anxiety: Relations to executive functioning in an at-risk sample. *Cognition and Emotion, 28*(5), 845–855.

https://doi.org/10.1080/02699931.2013.855173

Vandierendonck A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Van Breukelen, G. J. P. (2005). Psychometric modeling of response speed and accuracy with mixed and conditional regression. *Psychometrika, 70*(2), 359–376. https://doi.org/10.1007/s11336-003-1078-0

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Verhagen, J., Mulder, H., & Leseman, P. P. M. (2017). Effects of home language environment on inhibitory control in bilingual three-year-old children. *Bilingualism: Language and Cognition, 20*(1), 114–127. https://doi.org/10.1017/S1366728915000590

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology, 35*(1), 205–213. https://doi.org/10.1037/0012-1649.35.1.205

Willoughby, M. T., & Blair, C. B. (2016). Longitudinal measurement of executive function in preschoolers. In J. A. Griffin, P. McCardle, & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 91–113). American Psychological Association. https://doi.org/10.1037/14797-005

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2011). Contributions of modern measurement theory to measuring executive function in early childhood: An empirical demonstration. *Journal of Experimental Child Psychology, 108*(3), 414–435. https://doi.org/10.1016/j.jecp.2010.04.007

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment, 24*(2), 418–431. https://doi.org/10.1037/a0025779

Wilson, M., & De Boeck, P. (2004). Descriptive and explanatory item response models. In P. De Boeck & M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 43-74). Springer.

Yücel, M., Fornito, A., Youssef, G., Dwyer, D., Whittle, S., Wood, S. J., Lubman, D. I., Simmons, J., Pantelis, C., & Allen, N. B. (2012). Inhibitory control in young adolescents: The role of sex, intelligence, and temperament. *Neuropsychology, 26*(3), 347-356. https://doi.org/10.1037/a0027693

Zhao, X., Chen, L., & Maes, J. H. R. (2018). Training and transfer effects of response inhibition training in children and adults. *Developmental Science, 21*(1), 1–12. https://doi.org/10.1111/desc.12511

# Chapter IV

# Study 2

# Assessing Inhibitory Control in Kindergarten Children: Validity of Integrating Response Accuracy and Response Latency

A version of this chapter was published in:

Schulz, D., Segerer, R., Lenhard, W., Mangold, M., Schindler, J. & Richter, T. (2023). Assessing inhibitory control in kindergarten children: Validity of integrating response accuracy and response latency. *Cognitive Development, 68*, Article 101392. https://doi.org/10.1016/j.cogdev.2023.101392

Online supplementary materials, Data, and R code for the analyses reported in the present paper are available at the repository of the Open Science Framework (https://osf.io/t83yh/?view_only=d01a9610a6ba40eaa3c231734cb99400).

**Abstract**

Assessing inhibitory control in young children poses a challenge because of its rapid and non-linear development. This study examined the validity of integrating response accuracy and latency through a two-factor model, based on the data of 271 children who completed a computerized inhibitory control task. Although integrating response accuracy and latency slightly improved measurement precision, multigroup analyses of younger and older children showed inconsistent associations between response accuracy and latency if response latencies from incorrect responses were not excluded. A time-on-task analysis revealed that the extent of the accuracy gain by taking more time depended on the individual's skill level. The validity of task performance as an indicator of inhibitory control was highest when response accuracy was the primary determinant of the inhibitory control score and response latency was only considered after the child had surpassed an accuracy threshold to further improve the score. These findings suggest that integrating response accuracy and latency into a single score should only be performed for children who can maintain high accuracy levels despite giving fast responses.

*Keywords:* inhibitory control, executive function, structural equation modeling, generalized linear mixed model, assessment

## Assessing Inhibitory Control in Kindergarten Children: Validity of Integrating Response Accuracy and Response Latency

The ability to suppress inappropriate responses is essential for many cognitive activities (Diamond, 2013). Inhibitory control is also related to and predicts important facets of children's development (e.g., Ng et al., 2014; Rhoades et al., 2009). However, despite the importance of the construct in developmental and cognitive psychology, assessing inhibitory control remains a challenge, especially in young children and when focusing on a broader age range such as in longitudinal studies.

Inhibitory control is typically assessed with inhibition tasks such as variants of the Stroop tasks in which children have to suppress a predominant response to provide a correct response. For example, in the computerized Pointing Stroop Task, children are presented with an image of a dog and an image of a cat. In a first block, they must select the dog when they hear a bark and the cat when they hear a meow. In a second block, the inhibition block, the rule is reversed. They must click on the cat when they hear a bark and on the dog when they hear a meow (Berger et al., 2000; for an overview of measures, see Carlson, 2005; Garon et al., 2008). Inhibitory control measured with such tasks develops rapidly and in a non-linear fashion, which may complicate the assessment (e.g., Camerota et al., 2020). One way to improve the measurement of inhibitory control is to base the assessment not only on the accuracy of responses but also on the response latency. The present study examined the validity of model-based approaches to integrate accuracy and response latencies in assessments of inhibitory control. To that end, we analyzed the patterns in the association of response accuracy and latency as indicators of inhibitory control and the extent that these patterns change over the course of early childhood.

In the following sections, we first discuss inhibitory control in the context of executive functions (Diamond, 2013; Miyake & Friedman, 2012) and its development in young children. We describe how its development influences the measurement of inhibitory control and the challenges it creates. One approach to address these challenges is to combine the metrics of response accuracy and latency in inhibitory control tasks. In recent years, some researchers have proposed methods for integrating response accuracy and latency using formula-based approaches (Liesefeld & Janczyk, 2019; Vandierendonck, 2017) or model-based approaches (Camerota et al., 2020; Magnus et al., 2019). We discuss these approaches in the context of the developmental patterns of inhibitory control in young children. We specifically focus on the model-based approach proposed by Magnus et al. (2019; see also Camerota et al., 2020), which allows for a nuanced investigation of the association between response accuracy and latency in inhibitory control tasks by examining how item-level responses, both in terms of accuracy and latency, contribute to estimates of inhibitory control ability.

Our study applied this novel approach to analyzing data obtained from a well-established inhibitory control task - the computerized Pointing Stroop Task (Berger et al., 2000; Schulz et al., 2023). We examined the impact of including response latencies from incorrect responses on the association between accuracy and latency and how this association varies based on participants' inhibitory control ability. By investigating children's age as a potential moderator of the association between response accuracy and latency, we aimed to gain a more nuanced understanding of the association of response accuracy and latency when assessing inhibitory control in young children. Finally, we evaluated how different scoring methods to combine response accuracy and latency compare against two external criteria: fluid intelligence and children's age. This evaluation contributes to a better understanding of different methods of integrating response accuracy and latency in assessing inhibitory control in young children.

## Inhibitory Control

Inhibitory control as part of executive functioning is a core aspect of human cognition and development (Diamond, 2013; Miyake & Friedman, 2012). Together with working memory, inhibitory control may be regarded as a core factor of executive functioning (for an overview, see Müller & Kerns, 2015). Inhibitory control comprises interference control (inhibition of irrelevant stimulus information), cognitive inhibition (inhibition of irrelevant information in working memory), behavioral inhibition (inhibition of prepotent responses), and oculomotor inhibition (inhibition of reflexive eye movements; Nigg, 2000). Inhibitory control has been shown to be positively associated with academic success (Ng et al., 2014) and aspects of health behavior and outcomes (Allom et al., 2016; Mamrot & Hanć, 2019). In children, inhibitory control is closely related to socio-emotional competences (Rhoades et al., 2009) and to social, behavioral, and academic development (Domitrovich et al., 2017; Troller-Renfree et al., 2019). In the same vein, inhibitory control ability is negatively related to behavioral problems such as impulsivity and hyperactivity (e.g., Thorell et al., 2004).

Response inhibition specifically allows suppressing inappropriate responses such as those caused by learned or automated cognition and behavior (Verbruggen & Logan, 2008). The suppression of inappropriate responses, in turn, is an important factor in problem solving and therefore intelligence (Dempster, 1991). Inhibitory control further supports other executive functions in their functionality such as excluding irrelevant information from working memory (Bjorklund & Harnishfeger, 1990). These universal and children-specific effects and correlates of inhibitory control emphasize the importance of an accurate assessment of inhibitory control. However, the development of inhibitory control involves aspects that complicate its assessment, which we argue next.

**The Development and Assessment of Inhibitory Control in Young Children**

Assessing inhibitory control in young children is challenging for two reasons. First, children's inhibitory control ability improves quickly. Thus, tests designed for younger children are too easy for older children (Petersen et al., 2016), leading to ceiling effects in terms of response accuracy primarily at the age of 5 and older (Roebers, 2017). Second, and related to the first point, studies suggest that the development in inhibitory control is not linear and characterized by qualitative changes in the strategies children use to accomplish inhibitory control (Camerota et al., 2020; see also Chevalier et al., 2014).

### *Rapid Improvement and Ceiling Effects as Challenge for Assessing Inhibitory Control*

Given the rapid development of inhibitory control, tests designed for its assessment in children are mostly applicable within a relatively narrow age range. Petersen et al. (2016) concluded that most inhibitory control tests for children only yield informative response accuracy data for an age range of 3 years, depending on the target age. For example, the Shape Stroop Task (Kochanska et al., 1997) yields valid measurements of inhibitory control in children as young as 1.5 years but loses its validity after the age of 3.5. Other tests, such as the Simon Says Task (Strommen, 1973), validly measure inhibitory control until the age of 7 but only start to yield valuable information from the age of about 4.5 years. When a sufficient proportion of children reaches a ceiling effect, response accuracy no longer has discriminatory value. Using response latency in addition to accuracy seems to be a promising solution because children at all age levels exhibit large interindividual variance in how fast they respond in inhibition tasks (Jones, 2003). Therefore, measuring response latency when children reach a ceiling effect is a critical issue. Importantly, however, assessing only response latency is also insufficient because fast but inaccurate responses can hardly count as indicators of inhibitory control (Paap, 2019). In other words, a speed-accuracy tradeoff can occur (Luce, 1986) and must not be neglected in the interpretation of response latencies. The benefit of including response latency therefore depends on the ability level of the children taking the test.

Several formulae for integrating response accuracy and latency have been proposed to derive measures of efficiency (Liesefeld & Janczyk, 2019; Vandierendonck, 2017). For example, the *Inverse Efficiency Score* (Townsend & Ashby, 1983) combines aggregated response accuracy and latency scores into one score by dividing the mean response latency by the mean accuracy. The *Bin Score* approach, defined in Hughes et al. (2014), is a formula-based approach to control for response latencies in a control condition. The first step is to calculate the average response latency (RTc) across all participants and trials in the control condition. Next, the adjusted response latencies (RT-RTc) from correct trials in the experimental condition are sorted and divided into 10 deciles (or bins). For each participant, the number of correct trials in each bin is counted and assigned a weight based on the bin's position. Lastly, the number of error trials is also counted and assigned a weight of 20. The final Bin score is calculated by summing the weighted count of trials in each bin. Lower Bin scores indicate better performance.

Such integrated scores are useful and well interpretable in many contexts, especially when response accuracy is relatively high and shows little variability. However, most integrated scores rest on the assumption that response accuracy and latency information can be interpreted in the same way for all participants. In the context of developmental psychology, for example, this assumption is dubious when investigating the development of inhibitory control in young children. Zelazo et al. (2013) considered this limitation by assigning a score based on the proportion of correct responses first. Only if the accuracy was at or above the 80% threshold could the children's response latency improve the score further. This proposal is certainly an advancement compared to simple formula-based scores, but qualitative changes in the strategies used to complete inhibitory control tasks might require a more nuanced look into the association between response accuracy and latency.

### ***Qualitative Changes in Inhibitory Control Strategies as a Challenge for Assessing Inhibitory Control***

Camerota et al. (2020) found that for children who responded with low accuracy in the Hearts and Flowers Task (Davidson et al., 2006), which is a simple inhibition task, longer response latencies were associated with better executive function abilities, whereas for children who answered with high accuracy, shorter response latencies were associated with better executive function abilities. Chevalier et al. (2013; see also Chevalier et al., 2014; Chevalier et al., 2015) reported similar results in task-switching experiments. This developmental pattern has implications for investigating samples of younger children with a broad age range or for conducting longitudinal studies. As inhibitory control ability changes with age (Carlson & Moses, 2001; Kochanska et al., 1996; Roebers, 2017), inhibitory control can hardly be assessed with one integrative test score that fits all participants at all timepoints. In cross-sectional studies, younger children may not have reached a level of accuracy that allows for interpreting shorter response latencies as indicators of more efficient inhibitory control. Likewise, in longitudinal studies, the meaning of response latencies is likely to change when accuracy improves over the time points.

The *Dual Processing Theory* (Schneider & Fisk, 1983) provides an explanation for the changing association of response accuracy and latency in the development of young children. It differentiates between controlled and automated processing. Longer response latencies indicate an effective solution process when a task requires a high amount of control. Shorter response latencies indicate an efficient solution process when a task can be completed automatically. Whether a task can be completed in a controlled or automated fashion depends on both the task difficulty and the person's ability (Carlson et al., 1989). Inhibitory control tasks present a special case. They are designed not to be processed automatically because they aim for responses that go against strong prepotent tendencies, requiring conscious and active behavior regulation. However, with increasing mastery in inhibitory control, children become more efficient and are able to answer faster while maintaining high accuracy (cf. Dumont et al., 2022). Person ability varies greatly between and within age groups in children (Jones et al., 2003; Roebers, 2017).

**Assessing Inhibitory Control with a Model-Based Approach**

The discussion so far can be summed as follows: Only when children have reached a certain threshold of accuracy, faster response latencies can be seen as an indication of an efficient inhibitory control process, hence, indicating good inhibitory control. To examine this pattern in more detail, a model-based approach can be used such as structural equation modeling. This approach allows for a nuanced examination of how individual item-level responses (for accuracy and latency) contribute to the overall estimated inhibitory control ability.

The focus of previous studies utilizing structural equation models has been to integrate response accuracy and latency to achieve more precise measurement of inhibitory control (Camerota et al., 2020; Magnus et al., 2019). Magnus et al. (2019) explored the possibility of integrating response accuracy and latency into a structural equation model by using item response accuracies and latencies as indicators for inhibitory control. They proposed a bifactor model comprised of an *Inhibitory Control* (IC) factor and an orthogonal *Response Time* (RT) factor. The IC factor was based on item response accuracies and latencies. The RT factor was only based on the item response latencies and aimed to capture general response speed unrelated to inhibitory control. By constraining the factors to be orthogonal, they could separate the variance in response latency that provides information about IC ability from the variance reflecting the general speed of processing. Magnus et al. (2019) found that scores for inhibitory control based on the bifactor model were more precise than those based on the unidimensional model. Camerota et al. (2020) later expanded this approach. In a two-factor model, they included response latencies from a base condition (not requiring inhibition) as indicators for the RT factor to further control for general processing speed not influenced by inhibitory processes.

**Inclusion vs. Exclusion of Response Latencies from Incorrect Responses**

In many analyses of inhibitory control tasks with young children, response latencies for incorrect responses were excluded (e.g., Camerota et al., 2020; Davidson et al., 2006; Magnus et al., 2019; Ursache & Raver, 2014). The rationale is the uncertainty surrounding whether any inhibitory processes occurred in trials with incorrect responses. In such trials, children failed to inhibit the prepotent response. Therefore, including these data is believed to provide little meaningful insights into inhibition. However, excluding response latency data from incorrect responses may also lead to a loss of important information. Fast but inaccurate responses might indicate that the inhibitory control process was not initiated quickly enough, or possibly not initiated at all (Logan & Cowan, 1984). This suggests that these fast errors could provide meaningful information. Further, a more general association between response accuracy and latency often exists, for example, in the form of a speed-accuracy tradeoff (e.g., Wagenmakers et al., 2008). Additionally, younger children tend to make considerably more mistakes in inhibitory control tasks (Roebers, 2017), which means that removing response latencies for incorrect responses would result in a greater loss of response latency information from younger than from older children, a confound that could bias model estimates. As such, the inclusion of latencies from inaccurate responses, especially in comparison to a model in which these latencies are excluded, might provide a more comprehensive understanding of inhibitory control across different age groups. Therefore, it seems interesting to explore the empirical implications of including or excluding response latencies from incorrect responses when assessing young children's inhibitory control.

**Rationale of the Present Study**

The first aim of the present study was to thoroughly investigate the benefits of model-based integration of response accuracy and latency when assessing inhibitory control. A second aim was to shed light on the association of response accuracy and latency in inhibitory control tasks over the preschool years (3- to 6-year-olds). To that end, we examined data obtained with an established inhibitory control task, the computerized Pointing Stroop Task (Berger et al., 2000; Schulz et al., 2023).

To evaluate the potential benefits of combining item response accuracies and latencies in assessing inhibitory control, we applied a two-factor structural equation model similar to the one used by Magnus et al. (2019). However, we renamed the Response Speed factor to Response Time (RT) to better reflect the meaning of the factor. Longer response latencies result in higher RT loadings. We further extended this approach by also including item response latencies from congruent items as indicators for the RT factor (cf. Camerota et al., 2020). We hypothesized that the two-factor model including item response accuracies and latencies would show a more precise measurement of inhibitory control compared to a unidimensional model including only response accuracies from incongruent items as manifested in lower scale standard errors when estimating individual inhibitory control factor scores (Hypothesis 1).

The exclusion of response latencies for incorrect responses is a common approach in the research field. We explored whether and how loadings in the two-factor models differ with and without including response latencies for incorrect responses (Exploratory Research Question 1). The inclusion of response latencies for incorrect responses may introduce bias, particularly when comparing age groups. To examine the potential impact of this bias, we split the sample by median age into younger and older children and tested whether excluding or including response latencies from incorrect responses would affect comparability across the two groups (Exploratory Research Question 2). We tested whether the measurement models and factor loadings (i.e., the strength of the association between a latent variable and its indicators) were equal across younger and older children.

Both theory (Schneider & Fisk, 1983) and recent findings (Camerota et al., 2020) suggest that longer response latencies in younger children could be an indication of better inhibitory control because these children need to process the item in a controlled manner to reach the desired accuracy. Shorter response latencies in older children, who can process the items in a more efficient manner, should indicate better inhibitory control because they should be able, in principle, to respond fast and accurately. We expected these processing differences to show in the multigroup two-factor model when including response latencies for incorrect responses. Item response latencies of younger children should load positively on the Inhibitory Control factor because they rely more strongly on controlled cognitive processes (Hypothesis 2a), whereas item response latencies of older children should load negatively on the Inhibitory Control factor (Hypothesis 2b). An important note is that the validity of Hypotheses 2a and 2b is contingent on whether the measurement model is comparable across age group. If this is not the case, the interpretation of factor loadings is not possible.

We further aimed to explore the relationship between response accuracy and latency, with a particular focus on variations across participants' proficiency. To investigate this, we utilized the Generalized Linear Mixed Model framework (GLMM; De Boeck et al., 2011) at the item level. This approach allows assessing time-on-task effects, which can be understood as the individual's advantage from longer response latencies, and compare these with the person-specific intercept, which represents an individual's ability. Our objective was to improve our understanding of the complex interplay between the speed-accuracy tradeoff and individual differences. This tradeoff could become more apparent in less proficient children, who might show enhanced benefits from extended response latencies, perhaps due to their increased need for controlled processing. Conversely, we anticipated that proficient children could maintain accuracy with shorter response latencies, signaling cognitive efficiency.

Therefore, we hypothesized a negative correlation between the advantage from longer response latencies and person ability (Hypothesis 3).

Finally, we compared the model- and formula-based scoring approaches with regard to different indicators of validity (Exploratory Research Question 3). We examined the correlations among IC scores derived from our models, formula-based approaches, and the raw metrics of response accuracy and latency, to examine how strongly they converge. We also explored the associations of the various IC scores with age and a measure of fluid intelligence (a subscale of the SON-R 2 ½ – 7; Tellegen et al., 2006). This analysis enabled us to assess the external validity of the scoring approaches and evaluate their alignment with the construct of inhibitory control.

## Method

### Participants

Participants were 271 children (132 girls) from Switzerland and Germany recruited via childcare facilities, the research project's website, and local newspapers. Their mean age was 3.99 years (*SD* = 8.33 months; *Mdn* = 48 months) ranging from 2.83 to 5.25 years. Sociodemographic data were collected via a parent's questionnaire. The sample included 91 bilingual children. Of the 271 children, 29 spoke standard German, 74 spoke Bernese German, 40 spoke Basel German, 73 spoke French, 25 spoke Italian, and 30 spoke Turkish as their dominant language. The average parental education, measured by the highest education level of any parent, was high. The parents of three children did not complete the parental survey. From the remaining 268 children, 212 (79%) had at least one parent who held a university degree.

### Procedure

Data were collected in the context of a cross-national project that examined the cognitive, metacognitive, linguistic, and socio-emotional development of mono- and bilingual

3- to 6-year-olds. The children were tested in their childcare facilities, at home, or in dedicated lab spaces at universities. The sessions were conducted in individual settings with the research assistant and the child. The presence of a parent or educator during testing was permitted only if deemed necessary, as outlined in more detail in the online supplementary materials (p. 1). Testing took part at two different days for each child with 40 to 60 min sessions per day. Most of the testing, including the assessment of inhibitory control, was conducted on a 14 in. Windows convertible. The children were led through the tasks by a crocodile named Sammy who invited them on a virtual treasure hunt.

## Measures

To make our measures as language-fair as possible, story and instructions were presented in the language in which the tested child was most proficient. Language proficiency was inquired from the parents via a questionnaire and a telephone interview. If the research assistant observed during testing that the child was more proficient in the other language, the test was adapted, and instructions and stories were presented in the other language.

### *Inhibitory Control*

We used a shortened version of the computerized Pointing Stroop Task (cPST; Berger et al., 2000; Schulz et al., 2022) for the assessment of inhibitory control . The task was part of the test battery used in the project and were presented at different time points during the study, the order of which was counterbalanced across participants. During practice trials, if a child provided an incorrect answer, the trial was repeated. The children were instructed to respond quickly. The shortened version of the computerized Pointing Stroop Task took 5 min on average. As part of the story, Sammy and the child met a dolphin that was occupied with learning what dogs and cats sound like. The task consisted of two blocks. In both blocks, children were presented with an image of a cat and an image of a dog next to each other (position randomized between trials), and they simultaneously heard a bark or a meow. In the first block (congruent block), children were instructed to press the image matching the animal sound, that is, the cat when hearing a meow and the dog when hearing a bark. In the second block (incongruent block), children were instructed to press the image that contradicted the animal sound, that is, the dog when hearing a meow and the cat when hearing a bark. Each block consisted of eight trials. Before each block, two practice trials needed to be performed correctly by the child.

### ***Fluid Intelligence***

Fluid intelligence was assessed using the *Categories* subscale of the Snijders-Oomen Non-Verbal Intelligence Test (SON-R 2 ½ – 7; Tellegen et al., 2006). This subscale is designed to evaluate abstract thinking capability in young children. The assessment was administered at different time points depending on the randomization. The task consisted of two blocks. In the first block, children were presented with four to six cards depicting objects and were instructed to sort them into predefined categories. The second block required children to select one card from a pool of five cards, with the goal of choosing the card that shared a common feature with three other predefined cards.

## **Statistical Analysis**

Data from 10 children were excluded, resulting in a final sample of 261 children. Further details regarding the criteria for data exclusion and the number of children excluded for each task can be found in the online supplementary materials (p. 1). In line with Magnus et al. (2019), we excluded trials with response latencies less than 300 ms. Additionally, trials with response latencies more than 20 sec were excluded. Response latencies were logarithmically transformed. To enable comparisons between younger and older children, we divided our sample based on the median age of 48 months, yielding two subsamples: younger children ($n$ = 135, $M$ = 40.81 months, $SD$ = 3.55) and older children ($n$ = 126, $M$ = 55.57 months, $SD$ = 4.14). No significant performance differences were detected between monolingual and bilingual children. Additional analyses and results of this comparison are available in the online supplementary materials. All analyses were conducted using the R statistical software (R Core Team, 2022). Significance tests were based on a Type I error probability of .05.

***Confirmatory Factor Analysis***

We used confirmatory factor analysis for the item-level factor analysis models. All models were estimated with the full information maximum-likelihood estimator. Models including response latencies were estimated twice; once with response latencies for incorrect responses excluded, analogous to Magnus et al., (2019), and once using the data set with complete response latency data (Exploratory Research Question 1). The models were estimated with the package lavaan (Rosseel, 2012).

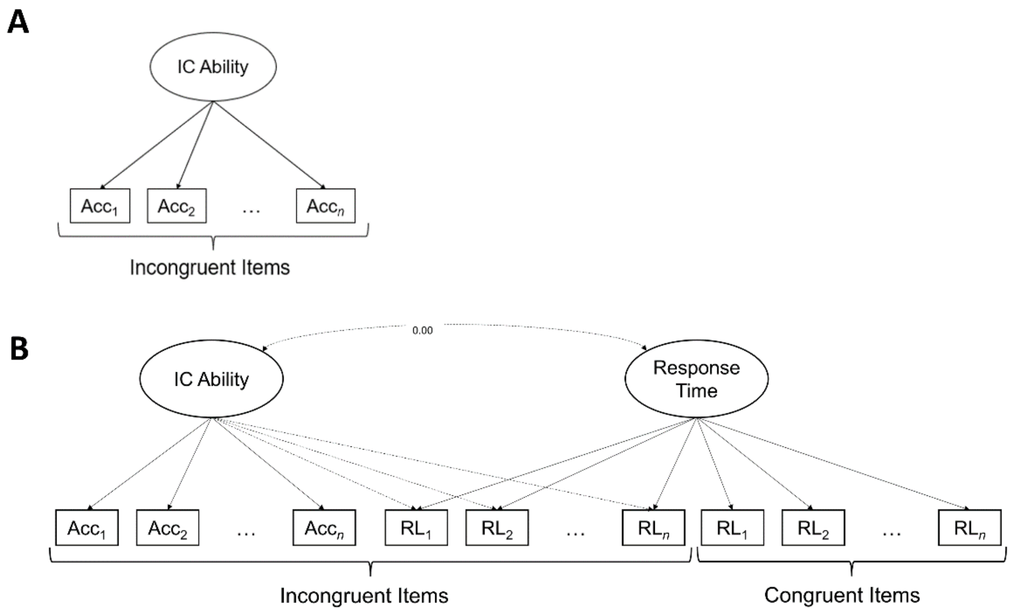
We first estimated a unidimensional model, using only item accuracies from incongruent items as indicators for the Inhibitory Control (IC) factor, as depicted in Figure 1, Panel A. The IC factor's mean was fixed to 0 and its variance to 1. Second, we estimated a two-factor model consisting of an Inhibitory Control (IC) and Response Time (RT) factor. Item response accuracies and item response latencies from incongruent items were used as indicators for the IC factor. For the RT factor, item response latencies from congruent and incongruent items were used as indicators. The IC and RT factors were constrained to be orthogonal (Figure 1, Panel B). For both factors, the mean was fixed to 0 and variance to 1. Unlike Magnus et al. (2019), we also evaluated the fit of the two-factor models using inferential statistics and descriptive goodness-of-fit measures (based on recommendations by Schermelleh-Engel et al., 2003). Measurement precision was determined by calculating the standard error of the factor scores.

Post-hoc, we investigated two other models detailed in the supplementary materials (Figure S2). Our aim was to better understand the factors influencing changes in measurement precision and model fit. One model expanded the unidimensional approach by including response latencies, while the second was a two-factor model accounting for item position by including error correlations between adjacent items. Through estimating these additional models, we sought to comprehend the impacts of different parameters on our model outcomes.

**Measurement Invariance Analysis.** We tested measurement invariance across age, that is, we examined whether and to what extent the measurement properties of the model are comparable across different age groups (Exploratory Research Question 2). We estimated a multigroup two-factor model with the structure described in the previous section for the groups of younger and older children, split based on median age. We first tested for configural invariance between the two groups. Configural invariance is established when the same underlying factor structure exists across different groups (Steinmetz et al., 2009). Given configural non-invariance, further comparisons between the groups regarding the measurements would not be valid. The lack of measurement model equivalency would preclude comparing the latent variable scores or the association between the latent variable and its indicators. If configural invariance is established, it is possible to examine whether excluding or including response latencies for incorrect responses affects metric invariance between the two groups. Metric invariance as a statistical concept enables testing whether the factor loadings or the strength of the association between a latent variable and its indicators are equal across different groups (Hypothesis 2). In addition, we carried out a correlational analysis of the associations of the average response accuracy, the average response latency, and the average response latency based on correct responses only (see supplementary results, p. 9). We evaluated these scores' correlations across the entire sample as well as within each age group. Our objective was to gain insights into the potential implications of including or excluding response latencies from incorrect responses.

**Figure 1**

*Defined Path Models for the Confirmatory Factor Analyses*




*Note.* Panel A: Unidimensional Model including only response accuracies. Panel B: Two-Factor Model. IC Ability: Inhibitory Control Factor. Response Time: Response Time Factor. $Acc_n$ = Item accuracy of item *n*. $RL_n$ = Response latency of item *n*.

### *Time-on-Task Analysis*

To investigate the impact of person ability on the association between response latency and accuracy (time-on-task effect; Hypothesis 3) and further explore the complexities of the speed-accuracy tradeoff, we used the GLMM framework (Goldhammer et al., 2014; Naumann & Goldhammer, 2017). We refer to response latency as “time-on-task” in accordance with the dual processing theory framework. GLMMs were estimated with the R packages lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017).

The original model for analyzing time-on-task effects models the solution probability as a function of person and task characteristics. For inhibitory control tasks, incorporating item effects into the model is less applicable due to near-identical trials in the incongruent block. The inclusion of item effects is justifiable only if it significantly improves variance explanation. Otherwise, we employ the reduced time-on-task model as presented in Formula 1:

$$ln\left(\frac{p_{pu}}{1-p_{pu}}\right) = \beta_0 + b_{0p} + (\beta_1 + b_{1p}) \cdot (time\ on\ task) + \beta_2 \cdot (base\ response\ speed) + b_{0u} \quad (1)$$

In this model, the probability of a correct solution ($p_{ipu}$) is modeled as a combination of the fixed intercept ($\beta_0$), fixed time-on-task effect ($\beta_1$), and the person-specific intercept and time-on-task effect. The person-specific intercept ($b_{0p}$) specifies the person ability. The higher the intercept, the higher the probability of that person performing an item correctly. The random time-on-task effect for person ($b_{1p}$) describes the person-specific time-on-task effect on the probability for a correct response. The higher the person-specific effect, the more a person benefits from taking their time for responding to the items. Performance in inhibitory control tasks must also be corrected for basic response speed. Therefore, we added a fixed effect of basic response speed ($\beta_2$), that is, the average log-response latency in congruent items ($z$-standardized; $\beta_2$).

***Correlational Analysis of Inhibitory Control Metrics and Covariates***

We conducted an additional analysis to correlate the scoring strategies based on our models and formula-based methods, in conjunction with raw response accuracy and latency metrics (Exploratory Research Question 3). To investigate the external validity of these scoring techniques and their alignment with the concept of inhibitory control, we further examined the correlation of these scores with age (in months) and the raw scores of our fluid intelligence measure.

**Results**

We report the model fit of the unidimensional and two-factor structural equation models and the impact of including response latencies on measurement precision (Hypothesis 1). We also report the impact of including response latencies for incorrect responses for the two-factor model for the overall sample (Exploratory Research Question 1) and for multigroup models by age (younger vs. older; Exploratory Research Question 2). We report on differences in factor loadings between younger and older children in the two-factor models but only if configural invariance was established (Hypothesis 2). Finally, we report the association between inhibitory control ability (i.e., their ability to correctly respond to items from the inhibition task) and the person specific time-on-task effect (i.e., how much children's accurate responding benefitted from longer response latencies) (Hypothesis 3).

**Improved Measurement of Inhibitory Control by Including Response Latencies: Two-Factor vs. Unidimensional Model**

All model fit indices and scale standard errors are provided in Table 1. The model that only included response accuracies from incongruent items (Figure 1, Panel A) had an acceptable model fit. The fit of the two-factor models (Figure 1, Panel B) was rather poor, regardless of whether response latencies from incorrect responses were included or excluded. The two-factor model showed a small improvement in measurement precision over the unidimensional model when only response accuracy was used (Hypothesis 1). The improvement was more pronounced when response latencies from incorrect responses were excluded. In a variant of the unidimensional model that incorporated both response accuracies and latencies from correct responses, a decreased standard error (0.25) was observed. However, this enhanced measurement precision came at the cost of a significantly deteriorated model fit (see supplementary material, Table S4). The analysis of a two-factor model accounting for error correlation between adjacent items showed an improvement in model fit (see supplementary materials; Table S4).

**Table 1**

*Model Fit Indices and Standard Errors of the Factor Scores*

| | Only Accuracy | Two-factor Model | |
|---|---|---|---|
| | | No error RL | Complete Data |
| RMSEA | .10 | .08 | .09 |
| CFI | .91 | .84 | .79 |
| TLI | .88 | .82 | .76 |
| *SE* | .40 | .34 | .38 |

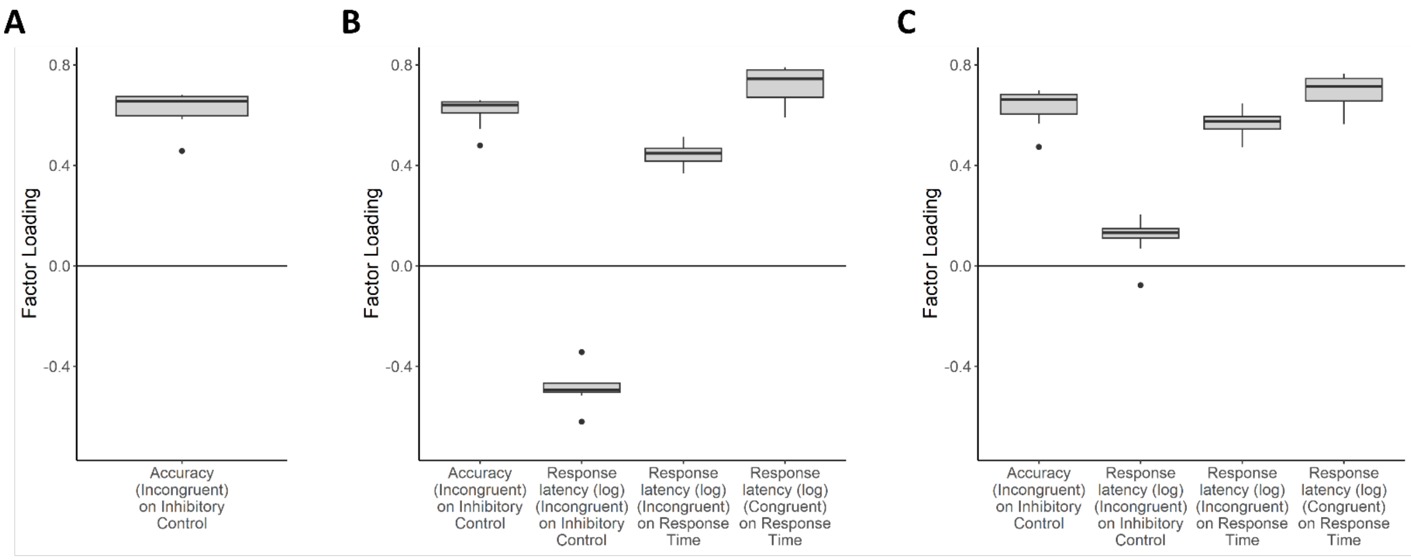
*Note.* No error RL = response latencies from incorrect responses excluded, Complete Data = response latencies from incorrect responses included, RMSEA = Root Mean Square Error of Approximation, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, *SE* = Standard Error of the factor scores.

**Effect of Excluding Response Latencies for Incorrect Responses**

Accuracies of all items loaded positively on the Inhibitory Control (IC) factor for the unidimensional and two-factor models (Figures 2, 3). Item response latencies loaded positively on the Response Time (RT) factor for all two-factor models (Figures 2, 3). Of primary interest were the loadings of item response latencies on the IC factor because an inclusion of response latencies for incorrect responses could introduce biases that influence the association between the IC factor and response latency (Exploratory Research Questions 1 and 2). Item metrics and response latency loadings from all two-factor models are provided in the online supplementary material (Table S2). Factor loadings are provided in Figure 2 for the unidimensional model (Panel A), the two-factor model estimated with the response latencies for incorrect responses excluded (Panel B) and the two-factor model estimated with complete response latency data (Panel C). The boxplots display the distributions of the standardized estimates for the loadings of item accuracies on the IC factor, the loadings of item response latencies on the IC factor, and the loadings of item response latencies on the RT factor.

**Figure 2**

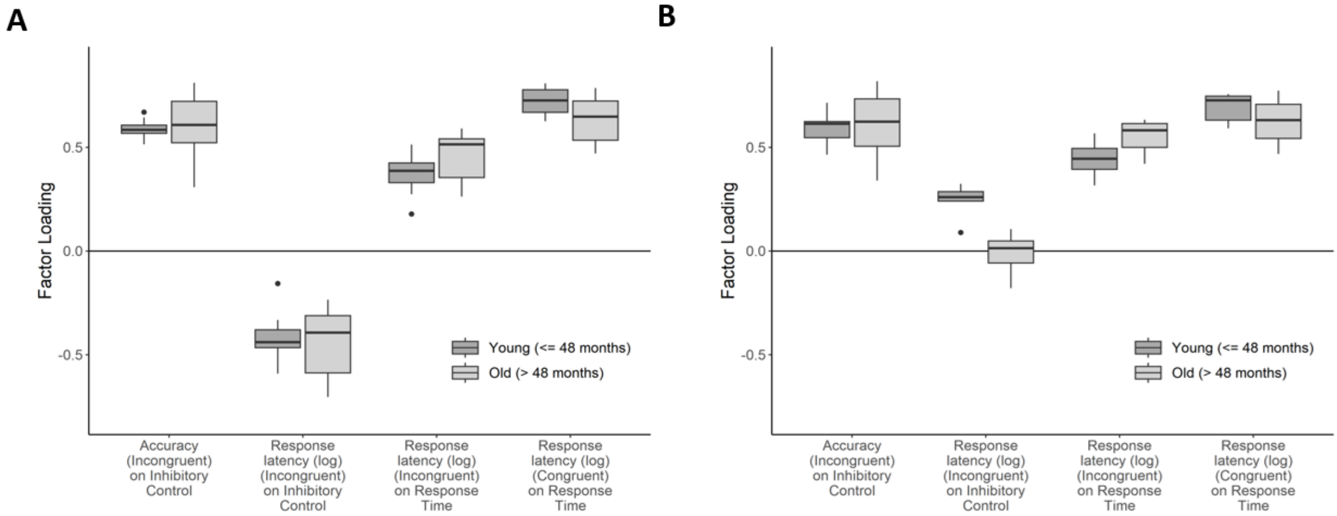
*Boxplots of Factor Loadings from the Unidimensional Model (Panel A) and the Two-Factor Models Without Response Latencies for Incorrect Responses (Panel B) and With Response Latencies for Incorrect Responses (Panel C)*




*Note.* The boxplots extend from the 25$^{th}$ to the 75$^{th}$ percentile with whiskers extending to most extreme value, maximally to the estimated 95% confidence interval. Panel A: Unidimensional model. Panel B: Two-factor model estimated without response latencies for incorrect responses. Panel C: Two-factor model estimated with response latencies from incorrect responses.

**Figure 3**

*Boxplots of Factor Loadings from the Multigroup Two-factor Models Without Response Latencies for Incorrect Responses (Panel A) and With Response Latencies for Incorrect Responses (Panel B) Separated by Age Group (Median Split)*




*Note.* The boxplots extend from the 25th to the 75th percentile with whiskers extending to most extreme value, maximally to the estimated 95% confidence interval. Panel A: Two-factor model estimated without response latencies for incorrect responses. Panel B: Two-factor model estimated with response latencies from all responses. Young = Participants at or below the age of 48 months. Old = Participants older than 48 months.

The loadings of item response latencies on the Inhibitory Control factor varied between the models estimated with and without response latencies for incorrect responses. When response latencies for incorrect responses were excluded, all factor loadings were significantly negative, whereas their inclusion resulted in predominantly positive loadings, four of which (out of eight) were significant (Exploratory Research Question 1). A substantial proportion of children made no errors (40%; additional information is provided in the online supplementary material, pp. 3ff). Nonetheless, younger children demonstrated lower accuracy

(67%) compared to older children (87%). Excluding response latencies for incorrect responses led to a greater loss of information for younger children than for older children. Thus, the pattern of missing data created by excluding latencies for incorrect responses systematically depended on the children's age.

**Multigroup Two-factor Models**

Complete measurement invariance information is provided in the online supplementary materials (Tables S3 and S5). We found that the model fit was rather poor before introducing constraints. The results of the invariance tests suggest configural non-invariance (Exploratory Research Question 2), both for the model without response latencies for incorrect responses ($\chi^2(488) = 1050.4$, $p < .001$, CFI = .789, RMSEA = .082) and for the model with complete response latency data ($\chi^2(488) = 918.3$, $p < .001$, CFI = .755, RMSEA = .094). This finding indicates that the measurements were not comparable across younger and older children in the current study, which suggests that the same underlying factor structure cannot be assumed for both groups. Consequently, when modeling inhibitory control by integrating response accuracy and latency based on the given model structure, comparisons between younger and older children's inhibitory control ability should be interpreted cautiously. Given that the validity of testing for metric invariance depends on configural invariance, caution should be exercised in interpreting the findings regarding metric invariance in the original models. Nevertheless, a visual inspection of the factor loadings from multigroup models suggests that the association between response latency and inhibitory control was similar when excluding response latencies for incorrect responses, but differed between younger and older children when complete response latency data were used (Hypothesis 2). In our examination of the best-fit model (post-hoc), which incorporated adjacent error correlation and excluded response latencies for incorrect, we observed an improvement in configural invariance. However, metric non-invariance persisted (see supplementary materials, Table S5).

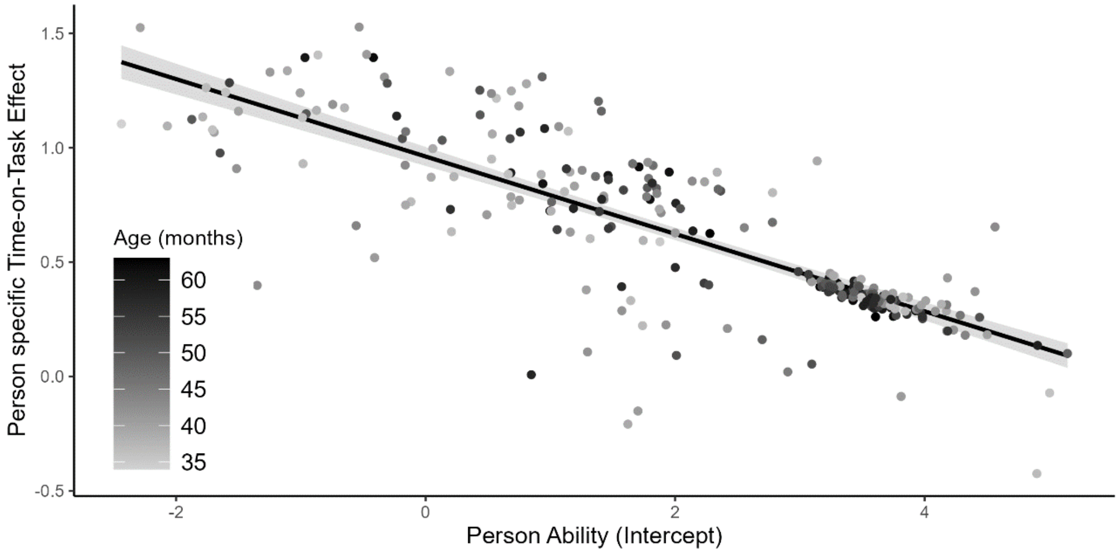
The correlation analysis in the supplementary materials (p. 9) emphasized the impact of including or excluding latencies for incorrect responses in younger vs. older children. When only correct response latencies were included, the correlation of response latency with response accuracy was significantly negative in both age groups. However, when all response latencies were included, a positive but insignificant correlation emerged for younger children, whereas for older children, the correlation remained significantly negative. Taken together, these results suggest that particularly for younger children, response latencies for incorrect responses may reflect different processes than response latencies for correct responses. The specific nature of this distinction could be elucidated by a detailed examination of the relationship between response latency and accuracy at the item level, which was done in the time-on-task analysis reported in the next section.

**Person-Specific Time-on-Task Effects**

As expected, the model including the fixed and person-specific time-on-task effect fit the data better than the null model ($\chi^2(4) = 101.02$, $p < .001$). The model fit remained the same when including item effects, ($\chi^2(2) = 0.00$, $p = 1.000$). The results revealed that the fixed time-on-task effect ($\beta_1$) was significant ($\beta = 0.55$; $z = 3.38$, $p < .001$), indicating that longer response latencies were positively associated with higher accuracy. In line with Hypothesis 3, we found a significant negative correlation between the person-specific intercept (i.e., person ability; $b_{0p}$) and the person-specific time-on-task effect ($b_{1p}$), $r(259) = -.81$, $p < .001$. This finding suggests that children who answered more accurately benefitted less from taking more time to answer to items. Furthermore, we found a significant and positive correlation between the person-specific intercept and age, $r(259) = .33$; $p < .001$, indicating that older children were less likely to make mistakes.

**Figure 4**

*The Association Between Person Ability (Random Intercept) and Person-Specific Time-On-Task Effect (Random Slope)*




*Note.* Each dot represents one participant. Person ability and person-specific time-on-task effect represent the random intercept and random slope from the GLMM (Formula 1) respectively. A regression line has been added with a 95% confidence interval. The cluster of cases on the right end of the regression line represents children who made no mistakes.

**Examination of Factor and Formula Based Scores**

The correlations between inhibitory control scores, fluid intelligence, and participants' age are provided in Table 2. Scores from the model with response latencies from incorrect responses excluded showed a slight advantage, exhibiting a more substantial correlation with response latency and age, alongside stronger associations with the Bin-O Score and Threshold Score. Moreover, aggregated response latencies from only correct responses were more strongly associated with age and fluid intelligence. Importantly, the Bin-O and Threshold Score demonstrated the highest correlations with both external criteria, age and fluid intelligence. The Bin-O Score adjusts for base response speed by controlling for the average response latency from the congruent condition across all participants. The Threshold Score crucially includes response latency (from the incongruent condition) but only after an accuracy threshold of 80%.

**Table 2**

*Correlations of Factor Scores, Formula Based Scores, Age, and Fluid Intelligence*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1. Two-Factor Model: Complete Data | | | | | | | | | |
| 2. Two-Factor Model: No Error RL | .92** | | | | | | | | |
| 3. Bin-Score | -.90** | -.95** | | | | | | | |
| 4. IES[a] | -.83** | -.80** | .76** | | | | | | |
| 5. Threshold Score[a] | .89** | .90** | -.92** | -.68** | | | | | |
| 6. Accuracy[a] | .99** | .95** | -.93** | -.84** | .91** | | | | |
| 7. RL (log): Full Data[a] | -.05 | -.30** | .41** | .16** | -.30** | -.10 | | | |
| 8. RL (log): No Error RL[a] | -.36** | -.65** | .68** | .38** | -.55** | -.42** | .91** | | |
| 9. Age (months) | .41** | .47** | -.54** | -.30** | .51** | .42** | -.45** | -.52** | |
| 10. Fluid Intelligence | .40** | .42** | -.48** | -.36** | .48** | .40** | -.33** | -.40** | .58** |

*Note.* Complete Data: All response latencies included. No Error RL: Response latencies from incorrect responses excluded. Bin-Score: As defined in Hughes et al. (2014); Lower scores indicate better performance. IES: Inverse Efficiency Score (Townsend & Ashby, 1983); Lower scores indicate better performance. Threshold Score: Score defined by Zelazo et al. (2013). Fluid Intelligence: Raw Scores from the SON-R. Score[a]: Score over incongruent items of the tests. * $p < .05$, ** $p < .01$.

## Discussion

In this study, we explored the possibility of improving the measurement of inhibitory control in young children by integrating response accuracy and latency in an inhibitory control task. Our first objective was to replicate the findings of Magnus et al. (2019). In particular, we aimed to replicate the improvement of measurement precision by incorporating response latencies and item-level response accuracies into a two-factor structural equation model over a unidimensional model that utilized only response accuracies. Consistent with the previous study, our results revealed a slight decrease in the standard error of the inhibitory control factor when including response latencies in the assessment of inhibitory control ability, indicating enhanced measurement precision. However, this decrease was very small and the model fit worsened after including response latencies. The diminished model fit of two-factor models after response latencies were included might be due, at least partly, to the different types of responses between the two measures, binary response accuracy vs. continuous response latency. Additionally, the association between response accuracy and latency may vary between individuals, which would suggest that a single model might not be suitable for estimating inhibitory control ability in a sample with a broad ability range. Post-hoc investigations revealed that integrating response latencies into a unidimensional model initially improved measurement precision but compromised the model fit, especially when latencies for incorrect responses were included. The model fit of the two-factor model improved when similarities between adjacent items were modelled via correlated errors.

In many studies investigating response latency from inhibitory control tasks, response latencies from incorrect responses were excluded (e.g., Camerota et al., 2020; Davidson et al., 2006; Magnus et al., 2019; Ursache & Raver, 2014). One reason for this methodological decision is the uncertainty of whether inhibitory processes take place in trials with incorrect responses. However, excluding response latency data from incorrect responses may lead to a loss of important information. Hasty, erroneous responses could indicate that inhibitory control processes were not activated at all or were initiated too slowly in children (Logan & Cowan, 1984). We examined the effect of including response latencies for incorrect responses. Inspecting the factor loadings showed that this decision had a significant influence on the model structure. Including all response latencies resulted in positive associations of longer response latencies with inhibitory control ability. Conversely, when response latencies from incorrect responses were excluded, shorter response latencies were associated with better inhibitory control ability. Crucially, the pattern of missing data in the model that excluded response latencies for incorrect responses was found to be highly dependent on participant age. Younger children made more errors, which lead to the exclusion of more data points. Thus, the data was not missing completely at random (Rubin, 1976), but contained a bias that might affect model estimations. Nevertheless, robust estimators such as the full information maximum likelihood estimator have proven effective in addressing missing at random data patterns (Enders, 2023; Enders & Bandalos, 2001).

We estimated a multigroup two-factor model with the same two-factor structure to investigate the differences between younger and older children. We found configural non-invariance, indicating that the measurement models were not comparable across age groups. An investigation of factor loadings in conjunction with the correlational analysis presented in the online supplementary materials (p. 9) suggests that the association between response accuracy and latency might vary across age groups when response latencies for incorrect responses were included. These findings suggest that longer response latencies from incorrect responses could potentially indicate enhanced inhibitory control, particularly in younger children who tend to make more mistakes. When we analyzed only response latencies for correct responses, shorter response latencies appeared to be associated with better inhibitory control ability across all children. Apparently, response latencies for incorrect answers provide unique information and reflects different processes than response latencies for correct responses. Until this distinction is fully understood, our advice, especially when combining response latency and accuracy in the assessment of young children, is to only use correct response latencies (besides response accuracy) for assessing inhibitory control. However, it is important to note that even using only correct response latencies lead to non-comparable measurement models, as indicated by the configural non-invariance, which makes it difficult to compare the scores of younger and older children.

In line with dual processing theory (Schneider & Fisk, 1983), it could be argued that for children still developing their cognitive abilities, long response latencies leading to correct responses may be indicative of inhibitory control proficiency (see also Goldhammer et al., 2014). Specifically, our results may suggest that taking the time necessary to suppress prepotent responses, particularly in younger children, might be one marker of inhibitory control ability, given it was associated with higher accuracy. Our results, derived from the time-on-task analysis, underscore that younger and less proficient children tend to gain more from taking additional time to respond accurately. However, the interpretation of these results needs careful consideration, particularly when examining the subset of children with the highest person ability. These children rarely make mistakes, which diminishes variance in response accuracy and thus the covariance with response latency. Nonetheless, the result effectively illustrates that the relationship between response latency and inhibitory control is contingent on the individual's ability to respond accurately. This finding underscores that proficiency in inhibitory control tasks is not merely about speed but is intrinsically linked to the accuracy of responses, particularly in developing populations.

These results align with the findings of Camerota et al. (2020) who observed that accuracy was negatively related to response latency for high-performing children, whereas for lower-performing children, accuracy was positively related to response latency. These results provide a clearer understanding of the observed shifts in factor loadings. Specifically, these shifts could be a consequence of the variation in the relationship between response accuracy and latency depending on age. Further, the findings are in line with research in other areas of cognitive development. In the area of cognitive flexibility, studies have demonstrated that younger children tend to have difficulty adapting to new rules and often continue to follow previously learned rules even when presented with a novel set of rules. In later phases of cognitive development, they acquire the ability to slow down their response latency, allowing them to adapt to new rules more effectively. This enhanced ability leads to a significant improvement in accuracy, which in turn can result in a constructive improvement in response latency. A recent longitudinal study by Dumont et al. (2022) investigated the development of cognitive flexibility in the Dimensional Change Card Sorting Task (Zelazo, 2006). The findings of the study revealed that the ability to slow down and accurately switch between rules at age 5 or 6 was a significant predictor of higher accuracy in the following year. Subsequently, higher accuracy at age 6 was found to be associated with shorter response latencies at age 7. Therefore, to exhibit cognitive flexibility, children must apparently first learn to slow down and switch rules accurately. In the area of self-regulation, Montroy et al. (2016) found that the development of self-regulation in children is best explained by a non-linear, exponential function, with children displaying different patterns of growth. Early developers showed a steep increase in self-regulation ability around the age of 4, whereas intermediate developers showed a less steep increase but with a slope that became steeper over time, speeding up around the age of 4.5. Late developers needed more time to develop, with their improvement in self-regulation starting later (around the age of 5) but still following an exponential function. These findings suggest that cognitive abilities, such as self-regulation, unfold at different rates among children, and that pronounced development occurs after reaching a certain age.

The assertion that faster response latencies are only a reliable indicator of inhibitory control if an accuracy threshold is met is further strengthened by comparing various methods of integrating response accuracy and latency to assess inhibitory control. Importantly, the connection with fluid intelligence was larger for scores based on the formula by Zelazo et al. (2013). This formula counts response latency only after a certain level of accuracy is reached. In simple terms, the faster a child can respond while staying accurate, the better their inhibitory control. Further, the correlational results with fluid intelligence suggest that the average response latencies for correct responses are a better indicator of inhibitory control than the average response latencies for all responses. These results underline that focusing on response latencies from correct answers may provide a more precise measure of inhibitory control in young children. Nonetheless, we suggest that longer latencies paired with lower accuracy could also offer valuable insights, as shown in the time-on-task analysis.

To enable a valid interpretation of response latencies in inhibitory control tasks, particularly in young children prone to errors, it seems necessary to establish an accuracy threshold that must be reached before response latencies can be used to augment the assessment of inhibitory control. Such a threshold model has been established in the development of basic reading processes (Juul et al. 2014; Karageorgos et al., 2019; Karageorgos et al., 2020). Almost all children can recognize words with high accuracy at the end of elementary school, but large individual differences remain in the degree of routinization. Based on longitudinal data, Karageorgos et al. (2020) found that word reading speed (or the efficiency of word reading) starts to develop only after children reach a certain accuracy in word reading tasks. Given the obvious similarities in the development of inhibitory control to basic reading processes (and other types of cognitive processes that become more routinized during development), the next logical step would be to use longitudinal data to examine whether such a threshold model also applies to the development of inhibitory control. In that case, the accuracy threshold would mark the point in individual development when the use of response latencies makes sense for the assessment of inhibitory control.

In addition to the empirical findings, understanding a critical conceptual distinction between response accuracy and response latency in inhibitory control tasks is essential. The primary objective of these tasks is to assess the ability to inhibit prepotent responses, with accuracy being the core measure and ultimate goal of the task. Achieving high accuracy indicates the ability to inhibit prepotent responses effectively. Therefore, prioritizing accuracy over speed is imperative and taking more time to answer accurately is desirable, as long as it leads to a substantial improvement. This perspective underscores the notion that faster response latencies should only be considered a positive sign of inhibitory control ability after a high level of accuracy has been established.

**Limitations**

Despite the consistent findings of this study, some limitations apply. First, the sample contained mostly children from families with a high socioeconomic status, which positively correlates with inhibitory control (Kałamała et al., 2020; Lipina et al., 2013). We expect more variability in the response accuracy distribution in samples of children from families with lower socioeconomic status. In the current sample, many children had already reached a ceiling in terms of response accuracy, as highlighted in the bar charts supplied in the online supplemental materials (Figure S1) and in the cluster of high ability children in the time-on-task analysis. Most likely, the associations found in the present study, especially in the time-on-task analysis, would have become even more pronounced in a sample with a higher variability of inhibitory control. The representativeness of our convenience sample cannot be guaranteed. However, the inclusion of participants with diverse cultural and national backgrounds added variability to our sample, potentially enhancing the external validity of our findings. Our cross-national recruitment of participants from Switzerland and Germany allowed for an examination of a range of cultural and national backgrounds in our analyses, although generalizability beyond these specific contexts may be limited. A second limitation is the close association of age and ability. The more important factor regarding the use of response latencies for assessing inhibitory control might be ability or more precisely the accuracy in inhibitory control tasks. Although age is a valid proxy for ability, assessing inhibitory control in younger children based on accuracy and older children based on response latency in inhibitory control tasks might be problematic because very proficient children will exist in the younger group, and children with poor inhibitory control will exist in the older group. Nonetheless, we chose to use age as the grouping variable for the multigroup models for two reasons. First, when splitting the children by median performance, very little variance in accuracy would be left in the high-performing group because many children had already reached a ceiling effect (information about the distribution of person-level accuracy are provided in the online supplementary materials, Table S1). This effect would have resulted in response accuracy not being a good indicator of inhibitory control because of a lack of variance. Secondly, we aimed to investigate the theoretical question of whether inhibitory control tasks can be analyzed uniformly across different age groups. The determining factor for this question will ultimately be the children's abilities, but this question is also relevant for evaluating inhibitory control in broader age ranges and in longitudinal analyses.

## Concluding Remarks

Despite these shortcomings, the present study contributes to our knowledge about the roles that response accuracy and latency play in the assessment of inhibitory control in children from 3 to 6 years of age. Previous studies have approached this topic by investigating the shortcomings of using one metric (Petersen et al. 2016) or by excluding response latencies from incorrect responses (e.g., Camerota et al., 2020; Davidson et al., 2006; Magnus et al., 2019; Ursache & Raver, 2014). However, the extent that the associations between response accuracy and latency (and therefore the diagnostic meaning of response latency) changes with children's age and ability level is still an underexplored research question. The findings of this study provide valuable insight into the nuances of measuring inhibitory control in children, highlighting the need for careful consideration of individual differences and test complexity when integrating response accuracy and latency. The present study adds to the growing body of research suggesting that the use of faster response latencies as indicators of executive function ability is dependent on a variety of factors, including age (Dumont et al., 2022) and skill level (Camerota et al., 2020). Our findings suggest that incorporating response latencies into assessments becomes more reliable when children can achieve and sustain a high accuracy. Until such a level of accuracy is maintained, using response latencies for correct responses could be a preferable alternative.

**References**

Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review, 10*(2), 168–186. https://doi.org/10.1080/17437199.2015.1051078

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Berger, A., Jones, L., Rothbart, M. K., & Posner, M. I. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. https://doi.org/10.3758/BF03207798

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*(1), 48-71. https://doi.org/10.1016/0273-2297(90)90004-N

Camerota, M., Willoughby, M. T., Magnus, B. E., & Blair, C. B. (2020). Leveraging item accuracy and reaction time to improve measurement of child executive function ability. *Psychological Assessment, 32*(12), 1118–1132. https://doi.org/10.1037/pas0000953

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Component fluency in a problem-solving context. *Human Factors, 31*(5), 489–502. https://doi.org/10.1177/001872088903100501

Chevalier, N., Huber, K. L., Wiebe, S. A., & Espy, K. A. (2013). Qualitative change in executive control during childhood and adulthood. *Cognition, 128*(1), 1–12. https://doi.org/10.1016/j.cognition.2013.02.012

Chevalier, N., James, T. D., Wiebe, S. A., Nelson, J. M., & Espy, K. A. (2014). Contribution of reactive and proactive control to children's working memory performance: Insight from item recall durations in response sequence planning. *Developmental Psychology, 50*(7), 1999–2008. https://doi.org/10.1037/a0036644

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience, 27*(6), 1125–1136. https://doi.org/10.1162/jocn_a_00782

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., & Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software, 39*(12), 1–28. https://doi.org/10.18637/jss.v039.i12

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157-173. https://doi.org/10.1016/0160-2896(91)90028-C

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development, 88*(2), 408–416.

https://doi.org/10.1111/cdev.12739

Dumont, É., Castellanos-Ryan, N., Parent, S., Jacques, S., Séguin, J. R., & Zelazo, P. D. (2022). Transactional longitudinal relations between accuracy and reaction time on a measure of cognitive flexibility at 5, 6, and 7 years of age. *Developmental Science, 25*(5), e13254. https://doi.org/10.1111/desc.13254

Enders, C. K. (2023). Fitting structural equation modeling with missing data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed., pp. 223-240). Guilford.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology, 106*(3), 608–626. https://doi.org/10.1037/a0034716

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*(3), 702–721. https://doi.org/10.3758/s13428-013-0411-5

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498–504. https://doi.org/10.1111/1467-7687.00307

Juul, H., Poulsen, M., & Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *Journal of Educational Psychology, 106*(4), 1096–1106. https://doi.org/10.1037/a0037100

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition, 204*, Article 104373. https://doi.org/10.1016/j.cognition.2020.104373

Karageorgos, P., Müller, B., & Richter, T. (2019). Modelling the relationship of accurate and fluent word recognition in primary school. *Learning and Individual Differences, 76*, 101779. https://doi.org/10.1016/j.lindif.2019.101779

Karageorgos, P., Richter, T., Haffmans, M.-B., Schindler, J., & Naumann, J. (2020). The role of word-recognition accuracy in the development of word-recognition speed and reading comprehension in primary school: A longitudinal examination. *Cognitive Development, 56*, Article 100949. https://doi.org/10.1016/j.cogdev.2020.100949

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263–277. https://doi.org/10.2307/1131849

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490-507. https://doi.org/10.2307/1131828

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Lipina, S., Segretin, S., Hermida, J., Prats, L., Fracchia, C., Camelo, J. L., & Colombo, J. (2013). Linking childhood poverty and cognition: Environmental mediators of non-verbal executive control in an Argentine sample. *Developmental Science, 16*(5), 697–707. https://doi.org/10.1111/desc.12080

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. Oxford University Press.

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*(7), 1296–1306. https://doi.org/10.1177/1073191117740953

Mamrot, P., & Hanć, T. (2019). The association of the executive functions with overweight and obesity indicators in children and adolescents: A literature review. *Neuroscience and Biobehavioral Reviews, 107*, 59–68. https://doi.org/10.1016/j.neubiorev.2019.08.021

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Montroy, J. J., Bowles, R. P., Skibbe, L. E., McClelland, M. M., & Morrison, F. J. (2016). The development of self-regulation across early childhood. *Developmental Psychology,* *52*(11), 1744–1762. https://doi.org/10.1037/dev0000159

Naumann, J., & Goldhammer, F. (2017). Time-on-task effects in digital reading are non-linear and moderated by persons' skills and tasks' demands. *Learning and Individual Differences, 53*, 1–16. https://doi.org/10.1016/j.lindif.2016.10.002

Ng, F. F.-Y., Tamis-LeMonda, C. S., Yoshikawa, H., & Sze, I. N.-L. (2014). Inhibitory control in preschool predicts early math skills in first grade: Evidence from an ethnically diverse sample. *International Journal of Behavioral Development, 39*(2), 139–149. https://doi.org/10.1177/0165025414538558

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Paap, K. R. (2019). The bilingual advantage debate: Quantity and quality of the evidence. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch34

Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., & Staples, A. D. (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review, 40*, 25–71. https://doi.org/10.1016/j.dr.2016.02.001

R Core Team (2022). *R: A language and environment for statistical computing* (Version 4.1.2) [Computer Software]. Retrieved from: https://www.R-project.org/

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social-emotional competence. *Journal of Applied Developmental Psychology, 30*(3), 310-320. https://doi.org/10.1016/j.appdev.2008.12.012

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Rosseel, Y. (2012). lavaan: An R Package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika, 63*, 581-590. https://doi.org/10.1093/biomet/63.3.581

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74.

Schneider, W., & Fisk, A. D. (1983). Attentional theory and mechanisms for skilled performance. In R. A. Magill (Ed.), *Memory and control of action* (pp. 119–143). North-Holland. https://doi.org/10.1016/ S0166-4115(08)61989-5

Schulz, D., Richter, T., Schindler, J., Lenhard, W., & Mangold, M. (2023). Using accuracy and response times to assess inhibitory control in kindergarten children: An analysis with explanatory item response models. *Journal of Cognition and Development, 24*(1), 82-104. https://doi.org/10.1080/15248372.2022.2119977

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality & Quantity: International Journal of Methodology, 43*(4), 599–616. https://doi.org/10.1007/s11135-007-9143-x

Strommen, E. A. (1973). Verbal self-regulation in a children's game: Impulsive errors on "Simon says". *Child Development, 44*(4), 849–853. https://doi.org/10.2307/1127737

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(10), 1231–1238. https://doi.org/10.1097/00004583-200210000-00013

Thorell, L. B., Bohlin, G., & Rydell, A.-M. (2004). Two types of inhibitory control: Predictive relations to social functioning. *International Journal of Behavioral Development, 28*(3), 193–203. https://doi.org/10.1080/01650250344000389

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press.

Troller-Renfree, S. V., Buzzell, G. A., Bowers, M. E., Salo, V. C., Forman-Alberti, A., Smith, E., Papp, L. J., McDermott, J. M., Pine, D. S., Henderson, H. A., & Fox, N. A. (2019). Development of inhibitory control during childhood and its relations to early temperament and later social anxiety: Unique insights provided by latent growth modeling and signal detection theory. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 60*(6), 622–629. https://doi.org/10.1111/jcpp.13025

Ursache, A., & Raver, C. C. (2014). Trait and state anxiety: Relations to executive functioning in an at-risk sample. *Cognition & Emotion, 28*(5), 845–855. https://doi.org/10.1080/02699931.2013.855173

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58*(1), 140–159. https://doi.org/10.1016/j.jml.2007.04.006

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., & 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

# Chapter V

# Study 3

# Balancing Accuracy and Speed in the Development of Inhibitory Control

A version of this chapter was submitted as:

Schulz, D., Lenhard, W., Mangold, M., Schindler, J. & Richter, T. (submitted). Balancing accuracy and speed in the development of inhibitory control. *Manuscript submitted for publication*,

Data and R code for the analyses reported in the present paper are available at the repository of the Open Science Framework (https://osf.io/2awp4/?view_only=1b110f7c5d764a28ae28f875214d20c2).

**Abstract**

Inhibitory control develops rapidly and non-linearly, making its accurate assessment challenging. This study aimed to investigate the non-linear relationship between accuracy and response latency in inhibitory control assessment of 3- to 6-year-old children in a longitudinal ($n$ = 431, 212 girls, $M$ = 4.86, $SD$ = 0.99) and a cross-sectional ($n$ = 135, 71 girls, $M$ = 4.24, $SD$ = 0.61) study. We employed a computerized Stroop Task to measure inhibitory control, with fluid intelligence serving as a covariate. Results from a growth curve analysis showed that children who reached an accuracy threshold of 80% earlier demonstrated faster improvements in response latency. Both the cross-sectional and the longitudinal findings demonstrated a positive association between response latency in the inhibitory control task and fluid intelligence but only when participants had achieved and maintained high accuracy. These results suggest that researchers should only consider response latency as an indicator of inhibitory control in children who manage to respond accurately in an inhibitory control task.

*Keywords:* inhibitory control, cognitive development, growth curve, response latency, response accuracy, fluid intelligence

## Balancing Accuracy and Speed in the Development of Inhibitory Control

Inhibitory control is a crucial cognitive ability that plays an essential role in children's cognitive and socio-emotional development (e.g., Rhoades et al., 2009). It is closely linked to the functioning of the prefrontal cortex and to fluid intelligence (Dempster, 1991) and is typically assessed with the Stroop paradigm or other tasks that require the inhibition of a prepotent response (e.g., Berger et al., 2000). However, the development of inhibitory control is not linear and characterized by qualitative changes in the strategies that children use to accomplish it. Rapid improvement in early childhood leads to ceiling effects of response accuracy, and response latency becomes more important and indicative of inhibitory control ability (Roebers, 2016). This non-linear development of response accuracy and latency in inhibitory control tasks leads to challenges when measuring inhibitory control in children because response latency can have different interpretations depending on the skill level of the child (Camerota et al., 2020).

In the following sections, we discuss inhibitory control, its close correlates such as fluid intelligence, and its non-linear development. In one longitudinal study and one cross-sectional study, we investigated the development of inhibitory control in young children and its association with fluid intelligence. We investigated whether response latency in an inhibitory control task develops differently between children who respond accurately and children who make more mistakes. We further examined the association between response latency in inhibitory control tasks and fluid intelligence depending on the level of response accuracy in cross-sectional and longitudinal designs. Understanding the association between response accuracy, latency, and inhibitory control is important for a valid assessment of inhibitory control in the preschool years and might help to better support children in their development of these important cognitive abilities.

**Inhibitory Control**

Inhibitory Control, as an aspect of executive functions, is a cognitive ability that is crucial for a wide range of cognitive activities (Diamond, 2013; Miyake & Friedman, 2012). As one central aspect, inhibitory control includes response inhibition (Harnishfeger, 1995). Response inhibition is the capacity to suppress inappropriate reactions, which allows the individual to direct their behavior in a way that conflicts with automated or learned behavior (Verbruggen & Logan, 2008). Suppressing such prepotent responses requires cognitive effort, which is commonly assumed to lead to increased error rates or response latencies when the required response is incongruent with an automatic response triggered by the stimulus (MacLeod, 1991). More accurate and faster responses in tasks that involve the suppression of inappropriate reactions are assumed to indicate better response inhibition.

The ability of inhibitory control is relevant for many cognitive activities and for performing overt actions. In children, inhibitory control has been linked to improved math skills (Ng et al., 2014), socio-emotional development (e.g., Rhoades et al., 2009) and decreased impulsivity and hyperactivity (Thorell et al., 2004). Moreover, inhibitory control can support other executive functions such as excluding irrelevant information from working memory (Bjorklund & Harshfeger, 1990), which can, in turn, contribute to fluid intelligence (Conway et al., 2003).

**Inhibitory Control and Intelligence**

Inhibitory control also seems to be relevant for tasks commonly used to measure fluid intelligence. Inhibitory control enables the suppression of irrelevant information, and this capability is necessary for problem-solving (Dempster, 1991) and is closely linked to the functioning of the prefrontal cortex (PFC; Moriguchi & Hikari, 2013). Studies have found that lesions in the PFC are negatively correlated with fluid intelligence (Roca et al., 2010). Furthermore, inhibitory control is important for disengaging from unsuccessful solutions

(Dempster, 1991), allowing for the flexibility which is necessary for many problem solving tasks.

Despite the conceptual and empirical linkages of inhibitory control and fluid intelligence, not all studies have yielded supporting evidence. Yücel et al. (2012) found a positive correlation between adolescents' intelligence and their response accuracy in a Stroop task. Friedman et al. (2006) also found a positive association between inhibitory control and multiple intelligence measures in young adults, but no correlation when both updating and inhibitory control were considered simultaneously as predictors (see also Duan et al., 2010). However, considering both response accuracy and latency instead of only one metric may provide a more comprehensive and valid assessment of inhibitory control (Camerota et al., 2019; Camerota et al., 2020; Magnus et al., 2019; Zelazo et al., 2013), which might lead to a stronger association with fluid intelligence (Schulz et al., 2023).

**The Development and Assessment of Inhibitory Control in Young Children**

Assessing inhibitory control in young children is challenging because of the rapid improvement in early childhood, which leads to ceiling effects in response accuracy, primarily at the age of 5 and older (Roebers, 2017). Additionally, studies suggest that the development of inhibitory control is not linear and characterized by qualitative changes in the strategies children use to accomplish inhibitory control (Camerota et al., 2020).

Given the rapid development of inhibitory control, tests designed for its assessment in children are mostly applicable within a relatively narrow age range when only response accuracy is considered (Petersen et al., 2016). As soon as a sufficient proportion of children reaches a ceiling effect, response accuracy no longer has discriminatory value. As a result, response latency could then be considered in addition to accuracy. However, even in older children, response latency should be interpreted carefully as fast but inaccurate responses can hardly be seen as indicators of inhibitory control. Generally, a speed-accuracy tradeoff (Luce, 1986) can occur in inhibitory control tasks.

Camerota et al. (2020) found that the association between response latency and response accuracy is not linear but depends on the level of response accuracy. In children who responded with lower accuracy, longer response latencies were associated with better executive function abilities, whereas for children who answered with high accuracy, shorter response latencies were associated with better executive function abilities. In children who are not yet able to answer with high accuracy, longer response latencies could indicate the ability to suppress a response long enough to answer correctly (c.f. Logan & Cowan, 1984). In contrast, shorter response latencies might indicate a more efficient inhibitory control in children who are able to respond with high accuracy. These children are able to maintain a high response accuracy despite increasing their response times.

The phenomenon of response latency being dependent on accuracy in children is not limited to inhibitory control tasks. In the literature on reading development, for example, studies have found that children must first reach an accuracy threshold before they can develop their response latency in a constructive manner (Juul et al., 2014; Karageorgos et al., 2019) and more quickly (Karageorgos et al., 2020).

This non-linear development is not only reflected in the response pattern of children when confronted with an inhibitory control task, but it is also supported indirectly by neurological differences between younger and older children. Specifically, younger children have been shown to have a more general neuronal activation than older children in inhibitory control tasks (Tamm et al., 2002). Older children, in contrast, show increased focal activation in the left inferior frontal gyrus, which has been found to be critical for inhibitory control (Swick et al., 2008). These results support the notion of a qualitative change in the strategies used to tackle inhibitory control tasks, given that differences in processing can be observed at the neuronal level.

In the context of inhibitory control, early development can be distinguished based on the level of accuracy achieved because accuracy serves as a more informative metric for younger children (Paap, 2019; Schulz et al., 2023). Zelazo et al. (2013) proposed an approach that emphasizes the importance of accuracy in assessing inhibitory control. This approach assigns a score based on the proportion of correct responses, prioritizing accuracy as the primary measure. Only when accuracy reaches or exceeds the threshold of 80%, response latency is considered as a contributing factor to the overall score. This recognizes that faster response latencies may not always indicate better inhibitory control. Instead, the attainment of the accuracy threshold serves as an indication that children have reached a sufficient level of inhibition. Faster response latencies prior to reaching this threshold might not reliably serve as indicators of inhibitory control because they can be associated with lower accuracy and suggest difficulties in inhibiting responses long enough to answer correctly (Heitz, 2014). However, once the accuracy threshold has been reached or surpassed, further development in inhibitory control is more likely to manifest in response speed, which appears to differentiate more prominently at higher levels of ability (Jones, 2003).

**Rationale of the Present Study**

Inhibitory control improves rapidly in young children (e.g., Kochanska et al., 1997). The rapid development leads to challenges in its assessment because the patterns reflecting the ability change with age and skill. For older children, who already have a better inhibitory control ability, response latency becomes a more important indicator of inhibitory control ability because response accuracy can easily reach a ceiling (Roebers, 2017). Older children are able to maintain a high accuracy even with shorter processing time. For younger children, who still struggle with inhibition, fast response latencies may not be indicative of inhibitory control. Faster response times may even signal the inability to inhibit an automated response long enough to respond according to the task instructions. In other words, a speed-accuracy tradeoff may occur in younger children (Heitz, 2014; Wagenmakers et al., 2008). We argue that response latency becomes indicative of inhibitory control ability only when children can reliably inhibit the prepotent response.

We expect children who are already able to answer accurately, that is, at or above a threshold of 80% (c.f., Zelazo et al., 2013), to show an accelerated improvement in response latency because these children already possess a sufficiently effective inhibitory control to become faster and therefore more efficient. Children who have not yet reached that accuracy threshold need more cognitive resources and thus more time to control their responses. In longitudinal Study 1, we investigated whether response latency in an inhibitory control task develops differently between children who are able to answer accurately (accuracy of at least 80%) and children who have not yet acquired this level of inhibition. We pursued this question in a longitudinal study with three timepoints in an age range from 3-6 years (Study 1). We expected children who reached the threshold early would display a steeper growth of response speed than children who reached the threshold later. This pattern of responses would mean a steeper decrease in response latency (Hypothesis 1).

Furthermore, response latency should be associated with close correlates of inhibitory control for children who are already able to inhibit the prepotent response reliably. We investigated this association cross-sectionally at Time 1 and Time 3 of our longitudinal study (Study 1) and with a cross-sectional study (Study 2). We expected to find an interaction of response latency and accuracy because response latency and fluid intelligence should be negatively related in children who reached the threshold, that is, faster responses should be associated with a higher intelligence in those children. In contrast, no such relationship between response latency and fluid intelligence was expected in children who had not (yet) reached the accuracy threshold. In sum, the expected pattern of the interaction was such that children who are both accurate and fast in the inhibition task will also show the highest intelligence scores (Hypothesis 2). We expected these results to occur cross-sectionally at Time 1 (Hypothesis 2a) and Time 3 (Hypothesis 2b) of Study 1 (the two timepoints at which intelligence was assessed) and also for the data from the cross-sectional Study 2 (Hypothesis 2c).

Apart from these cross-sectional results, we assumed that integrated inhibitory control ability (i.e., the interaction between accuracy threshold reached and response latency) at a younger age also has a predictive value for intelligence at a later age. In line with the cross-sectional hypotheses, we expected the interaction between reaching the accuracy threshold and response latency at Time 2 to significantly predict intelligence at Time 3. This interaction was expected to be due to a negative association between response latency at Time 2 and fluid intelligence at Time 3 in children who had reached the threshold whereas no such relationship was expected for children who had not (yet) reached the threshold at Time 2 (Hypothesis 3).

## Study 1

### Method

#### *Design and Procedure*

Data were collected in a cross-national longitudinal study examining cognitive, metacognitive, linguistic, and socio-emotional development of mono- and bilingual children. We aimed to test every child three times with 6- to 12-month intervals between the testing sessions. However, children were also admitted to the study if they were too old to take part at previous timepoints, which means that they entered the study at Time 2 or Time 3. Children at the age of 3;0 to 4;11 started at Time 1, 4;0- to 5;11-year-olds started at Time 2, and 6;0- to 6;5-year-olds started at Time 3. Whether 4-year-old children started at Time 2 depended on whether they could still be assessed two times in the scope of the project.

Children were assessed in their childcare facilities, at home, or in dedicated lab spaces at the universities involved in the study. The sessions were conducted in individual settings with the research assistant and the child. A parent or an educator was present when necessary. Most of the testing, including the assessment of inhibitory control, was conducted on a 14 in. Windows convertible. The children were led through the tasks by a crocodile named Sammy who invited them on a virtual treasure hunt during testing at Time 1 and Time 3 and who needed help to find home during testing at Time 2. For Time 1, testing took part on two different days for each child with 40 to 60 min sessions per day. Testing at Time 2 and Time 3 took part on one day in a 75 to 120 min session. Children received a small gift after every timepoint testing.

***Participants***

A total of 431 children (212 girls) from Germany and Switzerland participated in the study. Participant recruitment was conducted via childcare facilities, the website of the research project, and local newspapers. Overall, 264 children (251 after exclusion) were tested at the first and at least one other timepoint. Subsample sizes, dropout, and age information are provided in Table 1. Sociodemographic data were collected via a parent's questionnaire. Among the 431 children who participated in the study, 73 spoke High German, 120 spoke Bernese German, 147 spoke Basel German, and 147 spoke French as their primary societal language. The sample included 166 bilingual children, with 93 speaking Italian and 73 speaking Turkish as their heritage language. The average parental education, as measured by the highest education level of any parent, was high. Two children had incomplete parental surveys, leaving a total of 429 children. Among them, 294 (69%) had at least one parent with a university degree.

**Table 1**

*Subsample Sizes Over Timepoints*

| Measurement point | *n* | $n_{girls}$ | Dropout | New | First and two | Age *M* (*SD*) |
|---|---|---|---|---|---|---|
| t1 | 321 | 162 | 65 | 321 | 0 | 3.93 (0.68) |
| t2 | 354 | 178 | 94 | 98 | 256 | 5.04 (0.70) |
| t3 | 280 | 144 | 0 | 20 | 264 | 5.66 (0.71) |

*Note.* Eight children took part in Time 1 and Time 3 testing but not in Time 2 testing. Dropout = Children who did not partake in the next timepoint. New = Children who were entirely new to the study. First and two = Children who completed the first measurement point and at least one other. Age = Age in years.

### *Measures*

**Computerized Pointing-Stroop Task.** We used a shortened version of the computerized Pointing-Stroop Task (cPST; Berger et al., 2000; Schulz et al., 2023) for the assessment of inhibitory control, which was part of the test battery used in the project. The cPST took 5 min on average and was assessed at every timepoint. As part of the story, Sammy and the child met a dolphin that was occupied with learning what dogs and cats sound like. The task consisted of two blocks. In both blocks, children were presented with an image of a cat and an image of a dog next to each other (position randomized between trials), and they simultaneously heard a bark or a meow. In the first block (congruent block), children were instructed to press the image matching the animal sound, that is, the cat when hearing a meow and the dog when hearing a bark. In the second block (incongruent block), children were instructed to press the image contradicting the animal sound, that is, pressing the dog when hearing a meow and pressing the cat when hearing a bark. Each block consisted of eight trials. Thus, given the accuracy threshold of 80% used in this study, children were allowed to make no more than one mistake to meet the threshold. Before each block, two practice trials needed to be solved correctly by the child. When an incorrect answer was given, the practice trial was repeated. The story and instructions were presented in the language in which the tested child was most proficient.

**Fluid Intelligence.** The *Categories* subscale of the Snijders-Oomen Non-Verbal Intelligence Test (SON-R 2 ½ – 7; Tellegen et al., 2006) was used as an indicator of the children's non-verbal fluid intelligence and was assessed at Time 1 and Time 3. This subtest is designed to assess abstract thinking capability in young children and was presented offline (i.e., non-computerized) as part of the project's test battery. In the first block, children were instructed to sort four to six cards depicting objects into predefined categories. In the second block, children were instructed to select one card from a pool of five cards that shared a common feature with three other predefined cards. During testing at Time 3, the SON-R was conducted in an adaptive manner. Children's abilities were assessed in a routing block first, and then based on their performance, they were subsequently tested in a forward or backward manner.

***Data Analysis Strategy***

Timepoint data for a child were excluded from the analyses if attempts of keeping the child's attentional focus on the task failed, if the research assistant needed to support the child in answering items, or if the child failed to comprehend the task. Only items from the incongruent block were used in the analyses. Data from trials with a response latency over 20 seconds were deemed invalid and excluded from the analyses. Additionally, we excluded data from anticipatory responses with a response latency of less than 300 milliseconds from the analyses. Response latencies were logarithmically transformed. All analyses were performed with the statistical software R (R Core Team, 2022). Linear mixed models were estimated with the packages lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017). All significance tests were based on a Type I error probability of .05. We conducted one-tailed significance tests for regression slopes, based on directional hypotheses. Mean accuracies and the percentage of children reaching the accuracy threshold at each timepoint are provided in Table 2.

**Growth Curve Analysis.** We used linear mixed models (Baayen et al., 2008) to analyze whether the development of response latency differs between children who could answer accurately and children who could not and to account for the unbalanced repeated measures of the present study. The model for analyzing the development of response latency is described in Formula 1:

$$\begin{aligned}\log(Response\ latency) = \\ &(\beta_0 + b_{0p}) + \\ &(\beta_1 + b_{1p}) \cdot age + \ \beta_3 \cdot (age\ at\ threshold) + \\ &\beta_4 \cdot (age * age\ at\ threshold) + \ b_{0u}\end{aligned} \tag{1}$$

In this model, logarithmically transformed item level response latencies were modeled as a linear combination of the fixed intercept ($\beta_0$), fixed effect of age (in years; centered; $\beta_1$), fixed effect of age at threshold (i.e., the age at which the child had reached the threshold; in years; centered; $\beta_3$) and the fixed interaction term of age and age at threshold ($\beta_4$). The intercept ($b_{0p}$) and the slope of age ($b_{1p}$) were allowed to vary randomly between participants.

**Table 2**

*Accuracy Information by Measurement Point*

| Measurement Point | *n* | Mean Accuracy | Percent Reaching Threshold |
|---|---|---|---|
| Study 1 | | | |
| Time 1 | 304 | .77 | 40 % |
| Time 2 | 344 | .88 | 80 % |
| Time 3 | 280 | .95 | 93 % |
| Study 2 | | | |
| Time 1 | 125 | 0.82 | 0.67 % |

*Note*. Mean Accuracy = overall accuracy in incongruent items across all samples in each wave. Percent Reaching Threshold = percentage of participants who reached the accuracy threshold of 80% in incongruent items. Deviations from the sample sizes reported in Table 1 stem from exclusion of participants.

**Inhibitory Control and Intelligence.** For the analysis investigating the connection between inhibitory control and fluid intelligence, response latency values were averaged for each participant across all incongruent items. In a first step, fluid intelligence was modeled using a linear regression analysis with response latency (*z*-standardized), whether the accuracy threshold was reached (weighted effect-coded; 1 = threshold reached), and their interaction term as predictors. Age (in years; centered) was included as a covariate. In a second step, simple slope analyses were performed to examine the relationship between response latency and fluid intelligence for the groups who had reached the accuracy threshold and those who had not (Cohen et al., 2003; Richter, 2007). To conduct these analyses, we estimated the regression model two additional times. One model utilized dummy coding of the predictor "accuracy threshold reached", with the threshold-reached group coded as 0 and the threshold-not-reached group coded as 1. In this model, the coefficient for response latency describes the association between response latency and fluid intelligence for children who had already reached the accuracy threshold. The second model used dummy coding with the threshold-reached group coded as 1 and the threshold-not-reached group coded as 0. In this model, the coefficient for response latencies reflects the relationship between response latency and fluid intelligence for children who had not yet reached the accuracy threshold. The analyses were conducted cross-sectionally for Time 1 and Time 3, as well as longitudinally using response accuracy and latency data from Time 2 and fluid intelligence scores from Time 3.

Data and R-code for the analyses reported in the present paper are available at the repository of the Open Science Framework (https://osf.io/2awp4/?view_only=1b110f7c5d764a28ae28f875214d20c2).

**Results**

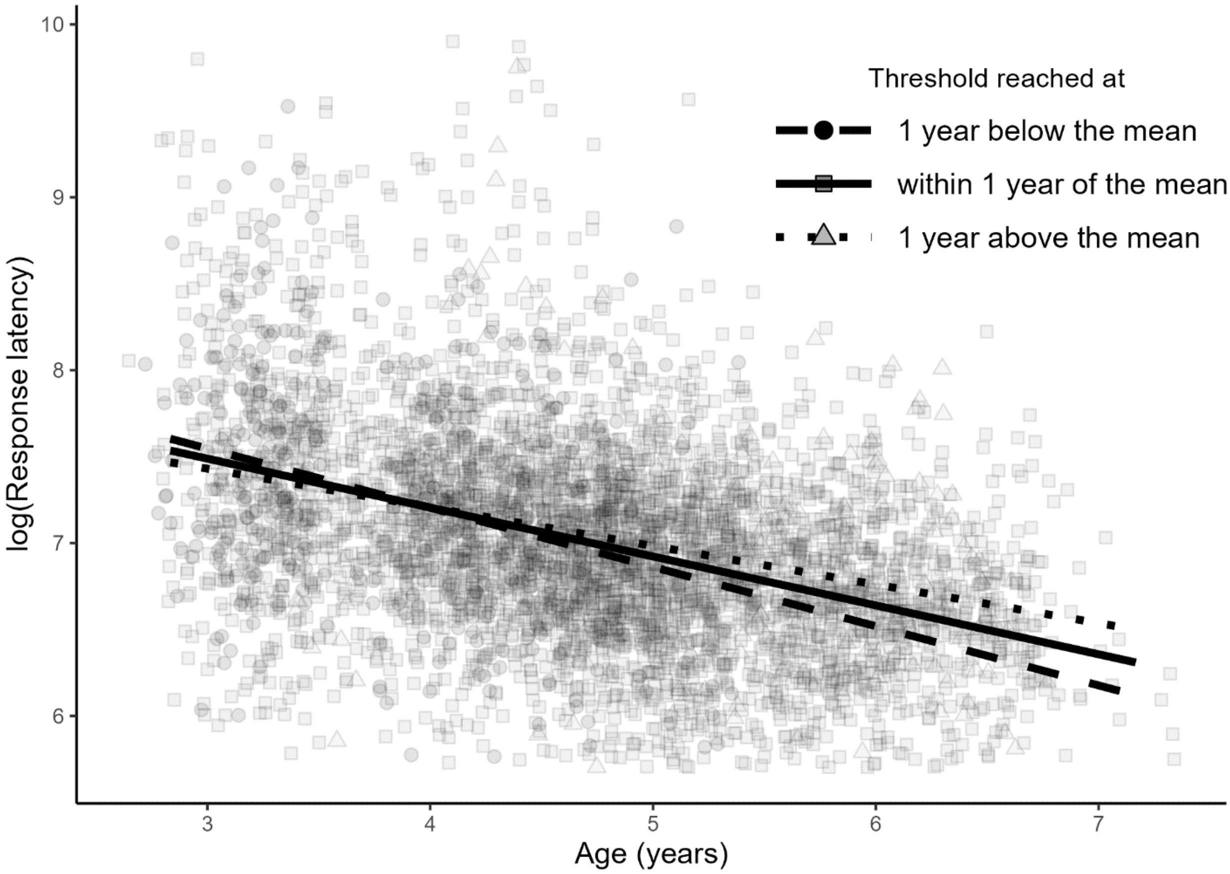
In 29 cases, timepoint data for children were excluded (21 at Time 1; 10 at Time 2; Table 1). The resulting sample consisted of 425 participants, as some children were only excluded at one of the timepoints. In the following sections, we first report results from the longitudinal growth curve analysis. Second, we report cross-sectional results concerning inhibitory control and fluid intelligence for Time 1 and Time 3. Lastly, we report the predictive value of inhibitory control ability as modeled by response accuracy and latency at Time 2 for intelligence at Time 3.

***Growth Curve Analysis***

We only included children who participated at Time 1 and at least one other timepoint ($n$ = 251) to ensure accurate assessment of their developmental trajectory. This approach also prevents potential underestimation of the age at which they reached the threshold for accuracy. Including the random effect of age significantly improved model fit over a model including only the linear term for age, age at threshold reached, and their interaction, $\chi^2(2) = 245.55$, $p < .001$. The main effect of age at threshold was not significant, β = 0.04; $t(227.04) = 1.43$, $p = .156$, whereas the main effect of the age term was significant, β = -0.28; $t(215.99) = -14.64$, *p* < .001. Children responded faster as they grew older. Most importantly, the interaction between the age term and the age at threshold was significant, β = 0.06; *t*(207.84) = -2.10, *p* = .037. In line with Hypothesis 1, children who reached the accuracy threshold earlier showed a steeper growth curve (Figure 1), that is, their response latency decreased faster.

**Figure 1**

*Impact of Age at Full Accuracy on Growth Curve of Response Latency*




*Note.* *n* = 57. Development of response latency (log-transformed) dependent on time (age in years). Divided into groups of children who reached the accuracy threshold earlier (1 year or more before the average age of reaching the threshold), children who reached the full accuracy threshold later (1 year or more after the average), and children who reached the threshold within one year of the average age at threshold reached.

***Inhibitory Control and Intelligence***

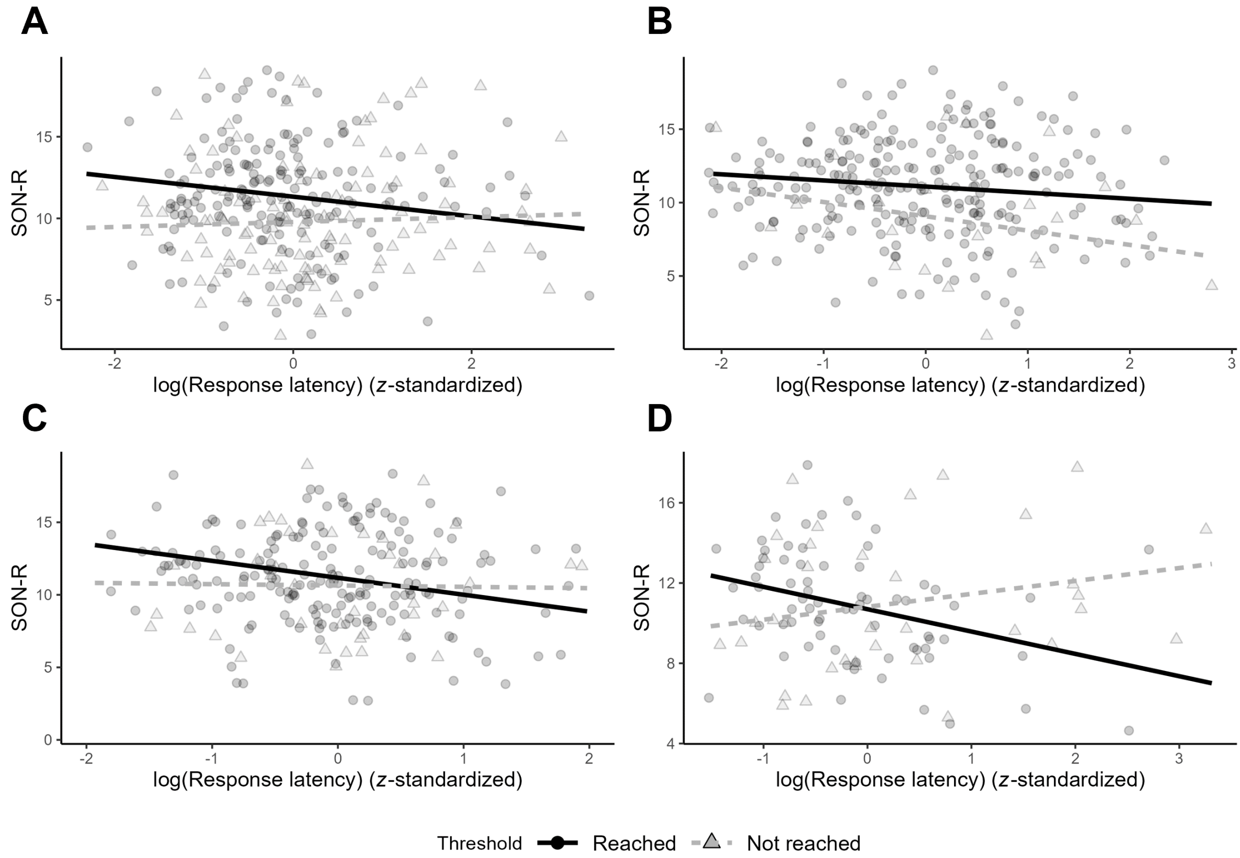
**Cross-sectional results.** At Time 1, the interaction between response latency and reaching the threshold was significant, β = -0.38; $t(287) = -1.81$, $p = .036$ (Figure 2, Panel A). In line with Hypothesis 2a, the relationship between response latency and fluid intelligence was significantly negative for children who had already reached the accuracy threshold, β = -0.60; $t(287) = -1.98$, $p = .024$, but was not significant for children who had not yet reached the threshold, β = 0.15; $t(287) = 0.50$, $p = .32$. The relationship between response latency and fluid intelligence was positive albeit nonsignificant for children who had not (yet) reached the accuracy threshold.

At Time 3, the interaction between response latency and reaching the threshold was not significant, β = 0.05; $t(259) = 0.88$, $p = .189$ (Figure 2, Panel B). However, in line with Hypothesis 2b, the relationship between response latency and fluid intelligence was significantly negative for children who had already reached the accuracy threshold, β = -0.42; $t(259) = -1.76$, $p = .040$, but nonsignificant for children who had not yet reached the threshold, β = -0.97; $t(259) = -1.64$, $p = 0.052$. The smaller $t$ value despite the larger estimate (β) resulted from a lower standard error for the predictor response latency in the group of children who had already reached the accuracy threshold ($SE = 0.24$) as opposed to the group of children who had not ($SE = 0.60$).

**Longitudinal results.** The interaction between response latency and reaching the accuracy threshold at Time 2 as a predictor for fluid intelligence at Time 3 was significant, β = -0.21; $t(235) = -2.27$, $p = .012$ (Figure 2, Panel C). In line with Hypothesis 3, the relationship between response latency and fluid intelligence was significantly negative for children who had already reached the accuracy threshold, β = -1.17; $t(235) = -4.17$, $p < .001$, but nonsignificant for children who had not yet reached the threshold, β = -0.09; $t(235) = -0.23$, $p = 0.410$.

**Figure 2**

*Relationship between Fluid Intelligence and Response Latency Moderated by Reaching the Accuracy Threshold: Cross-Sectional and Longitudinal Results*




*Note.* Panel A = Cross-sectional results of Time 1, Study 1. Panel B = Cross-sectional results of Time 3, Study 1. Panel C = Longitudinal results of Time 2 and Time 3, Study 1. Panel D = Cross-sectional results of Study 2.

# Study 2

## Method

### *Design and Procedure*

Cross-sectional data were collected as the pilot study for Study 1. Testing procedure was analogous to Time 1 of Study 1.

### *Participants*

Participants were 135 children (71 girls) from Switzerland and Germany recruited via childcare facilities, the experiment's website, and local newspapers. Their mean age was 4.24 years (*SD* = 7.42 months) ranging from 2.92 to 6.50 years. Sociodemographic data were collected via a parent questionnaire. Among the 431 children who participated in the study, 45 spoke High German, 34 spoke Bernese German, 27 spoke Basel German, and 29 spoke French as their primary societal language. The sample included 42 bilingual children, with 15 speaking Italian and 27 speaking Turkish as their native language. The average parental education as measured by the highest education level of any parent was high. Twenty children had incomplete parental surveys, leaving a total of 115 children. Among them, 88 (77%) had at least one parent with a university degree.

### *Measures and Strategy for Data Analysis*

The cPST was identical to the version used in Study 1 with the exception that every block consisted of four items, not eight as in Study 1. With a set accuracy threshold of 80%, children were not allowed to make mistakes to meet the threshold. The SON-R was identical with the version from Time 1 in Study 1. Analyses were identical to the cross-sectional analysis described in Study 1.

## Results

The data of 10 children were excluded from the analyses because they received help from the research assistant, parent, or educator, objectively were not able to comprehend the instructions (responded with two fingers simultaneously) or refused to carry out the task. For 18 children, SON-R testing was not conducted or not conducted properly, resulting in a sample of 107 children with cPST and SON-R values.

The interaction between response latency and reaching the threshold was significant, $\beta = -0.59$; $t(102) = 3.03$, $p = .002$ (Figure 2, Panel D). In line with Hypothesis 2c, the relationship between response latency and fluid intelligence was significantly negative for children who had already reached the accuracy threshold, $\beta = -1.11$; $t(102) = -2.55$, $p = .006$, but not significant for children who had not yet reached the threshold, $\beta = 0.64$; $t(102) = 1.62$, $p = 0.054$. Again, the relationship between response latency and fluid intelligence was positive albeit nonsignificant for children who had not (yet) reached the accuracy threshold.

## Discussion

The aim of the present study was to investigate the development of response latency in inhibitory control tasks among children who had already reached an accuracy threshold of 80% (c.f., Zelazo et al., 2013) and children who had not. The core hypothesis to be examined was that response latency can be considered a valid measure of inhibitory control only when children have reached a sufficient level of accuracy. Our findings from the growth curve analysis supported this hypothesis, showing that only after reaching this accuracy threshold, children are able to increase their efficiency. Importantly, our results demonstrated that children, who had already developed the ability to inhibit their responses and achieve accuracy, displayed a significantly steeper growth curve in response latency compared to those who had not yet reached this level of inhibition. These results indicate that longer response times early in the developmental process may even be a positive sign because taking time enables children to give accurate responses. Only when children have achieved a high level of accuracy, inhibitory control can become more efficient, which leads to speeded response latency.

The present study also examined the relationship between response latency and fluid intelligence, which is an important correlate of inhibitory control (e.g., Dempster, 1991). We hypothesized that response latency would be negatively associated with fluid intelligence in children who answered accurately, that is, children who answered faster would have a higher intelligence. However, we expected no association for children who answered inaccurately. The cross-sectional results from Study 1 and 2 revealed that response latency was negatively associated with fluid intelligence among children who answered accurately but not among children who answered inaccurately. The longitudinal results aligned with our hypothesis. Response latency in inhibition tasks at Time 2 was negatively associated with fluid intelligence at Time 3 only among accurately responding children. The lack of a significant interaction for Time 3 in Study 1 could be attributed to the overall high accuracy at that timepoint, implying that most participants had become proficient at the task. As a result, the group of children who had not reached the threshold was small and the variance in this predictor was low. Nevertheless, the significance of the simple slope only for children who had reached the accuracy threshold, as opposed to those who had not, in conjunction with a smaller standard error within the group that had reached this accuracy threshold, implies a more consistent and substantial negative association between response latency and fluid intelligence. Taken together, these findings indicate that shorter response latencies are a valid measure of inhibitory control only after children have achieved a sufficient level of accuracy.

The findings of the current study both align with and expand upon established and recent results and theories. In various domains, the relationship between response latency and skill may change depending on skill level. For example, Naumann and Goldhammer (2017) discovered that when poor readers took more time to work on a task, the probability of solving the task correctly was positively enhanced. This finding is consistent with the dual processing theory (Schneider & Chein, 2003), which posits that when individuals lack proficiency in a task, controlled processing is necessary. In such cases, longer response latencies are indicative of control. In research on cognitive flexibility, younger children tend to adhere to previously learned rules, without slowing down even when they are presented with a novel set of rules (Dumont et al., 2022). However, as children grow older and their cognitive development continues, they gradually acquire the capacity to slow down their response latency, which allows them to adapt to new rules (Zelazo et al., 2013). Dumont et al. (2022) conducted a longitudinal study that investigated the development of performance in the Dimensional Change Card Sorting Task (DCCS), an adapted version of Zelazo's (2006) task. According to their findings, demonstrating the ability to slow down and adjust to new rules in the task at around 5-6 years old was a significant predictor of improved accuracy the following year. Additionally, the results showed that higher accuracy at 6 years predicted shorter response latencies at 7 years. In other words, for children to exhibit cognitive flexibility, they must first learn to slow down and accurately switch rules. This development is deemed desirable because it leads to better accuracy. Only after this stage, faster response times indicate improved cognitive flexibility.

Recent findings on inhibitory control are in agreement with this pattern. For example, Camerota et al. (2020) observed that the connection between inhibitory control ability and response latency differed based on skill level. In the mixed condition of the Hearts and Flowers Task (Davidson et al., 2006), which requires not only cognitive flexibility but also inhibition, response latency was found to be negatively related to executive function ability for children who answered questions more accurately. This finding indicates that shorter response latencies were indicative of better performance. However, for children who answered questions less accurately, response latency was positively related to executive function ability, suggesting that longer response latencies were indicative of better performance. Together with the findings of the present study, response latency alone apparently is not a sufficient measure of inhibitory control in young children, and it should not be used as a standalone metric. However, response accuracy may lose its discriminatory value once children reach a ceiling effect (Petersen et al., 2016). Hence, studies with a broad age and skill range or longitudinal studies need to consider both response latency and accuracy to obtain a complete understanding of inhibitory control development.

Most scores integrating both metrics assume that response accuracy and latency information can be interpreted in the same way for all participants (for examples, see Liesefeld & Janczyk, 2019; Vandierendonck, 2017), which may not be the case in developmental psychology. Inhibitory control development is a complex process, and individual differences in the strategies used to complete inhibitory control tasks can lead to different associations between response latency and accuracy for different children. Therefore, an integration method that differentiates between children for whom faster response latencies are actually an indicator of inhibitory control would be more appropriate. Zelazo et al. (2013) proposed an approach that considers accuracy first and only considers response latency when accuracy is at or above 80%. Our results support using an approach that considers accuracy as the primary measure of inhibitory control and incorporates response latency only when accuracy is at or above a certain threshold when assessing inhibitory control in young children.

The insights provided by the present study notwithstanding, some limitations should be noted. One limitation of the present study is the focus on only one specific (albeit typical and valid; Schulz et al., 2023) inhibitory control task, and future research should explore the relationship between response latency and accuracy for different types of inhibitory control tasks, preferably measuring different kinds of inhibitory control (c.f. Nigg, 2000; Simpson &

Carroll, 2019). In conclusion, our research contributes to a better understanding of the development of inhibitory control and the relationship between response latency and accuracy in children, highlighting the importance of considering both measures when evaluating inhibitory control. The findings have significant implications for researchers and clinicians who use inhibitory control tasks to assess children's cognitive development. Considering response latency and accuracy when interpreting task performance is crucial, especially for children who are approaching a ceiling in response accuracy. An integration method that considers faster response latencies as indicators for inhibitory control only when accuracy can be maintained is best aligned with our results.

**References**

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Berger, A., Jones, L., Rothbart, M. K., & Posner, M. I. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. https://doi.org/10.3758/BF03207798

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*(1), 48-71. https://doi.org/10.1016/0273-2297(90)90004-N

Camerota, M., Willoughby, M. T., & Blair, C. B. (2019). Speed and accuracy on the Hearts and Flowers task interact to predict child outcomes. *Psychological Assessment, 31*(8), 995-1005. https://doi.org/10.1037/pas0000725

Camerota, M., Willoughby, M. T., Magnus, B. E., & Blair, C. B. (2020). Leveraging item accuracy and reaction time to improve measurement of child executive function ability. *Psychological Assessment, 32*(12), 1118–1132. https://doi.org/10.1037/pas0000953

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates Publishers.

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

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157-173. https://doi.org/10.1016/0160-2896(91)90028-C

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year-old children. *Psychological Test and Assessment Modeling, 52*(4), 419–431.

Dumont, É., Castellanos-Ryan, N., Parent, S., Jacques, S., Séguin, J. R., & Zelazo, P. D. (2022). Transactional longitudinal relations between accuracy and reaction time on a measure of cognitive flexibility at 5, 6, and 7 years of age. *Developmental Science, 25*(5), e13254. https://doi.org/10.1111/desc.13254

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). Academic Press. https://doi.org/10.1016/B978-012208930-5/50007-6

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*: 150. https://doi.org/10.3389/fnins.2014.00150

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498–504. https://doi.org/10.1111/1467-7687.00307

Juul, H., Poulsen, M., & Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *Journal of Educational Psychology, 106*(4), 1096–1106. https://doi.org/10.1037/a0037100

Karageorgos, P., Müller, B., & Richter, T. (2019). Modelling the relationship of accurate and fluent word recognition in primary school. *Learning and Individual Differences, 76*, 101779. https://doi.org/10.1016/j.lindif.2019.101779

Karageorgos, P., Richter, T., Haffmans, M.-B., Schindler, J., & Naumann, J. (2020). The role of word-recognition accuracy in the development of word-recognition speed and reading comprehension in primary school: A longitudinal examination. *Cognitive Development, 56*, Article 100949. https://doi.org/10.1016/j.cogdev.2020.100949

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263–277. https://doi.org/10.2307/1131849

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. Oxford University Press.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*(7), 1296–1306. https://doi.org/10.1177/1073191117740953

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Moriguchi, Y., & Hiraki, K. (2013). Prefrontal cortex and executive function in young children: A review of NIRS studies. *Frontiers in Human Neuroscience, 7*: 867. https://doi.org/10.3389/fnhum.2013.00867

Naumann, J., & Goldhammer, F. (2017). Time-on-task effects in digital reading are non-linear and moderated by persons' skills and tasks' demands. *Learning and Individual Differences, 53*, 1–16. https://doi.org/10.1016/j.lindif.2016.10.002

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Paap, K. R. (2019). The bilingual advantage debate: Quantity and quality of the evidence. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch34

Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., & Staples, A. D. (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review, 40*, 25–71. https://doi.org/10.1016/j.dr.2016.02.001

R Core Team (2022). *R: A language and environment for statistical computing* (Version 4.2.1) [Computer Software]. Retrieved from: https://www.R-project.org/

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social–emotional competence. *Journal of Applied Developmental Psychology, 30*(3), 310–320. https://doi.org/10.1016/j.appdev.2008.12.012

Richter, T. (2007). Wie analysiert man Interaktionen von metrischen und kategorialen Prädiktoren? Nicht mit Median-Splits! [How does one analyze interactions of metric and categorical predictors? Not with median splits!]. *Zeitschrift für Medienpsychologie, 19*(3), 116-125. https://doi.org/10.1026/1617-6383.19.3.116

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., & Duncan, J. (2010). Executive function and fluid intelligence after frontal lobe lesions. *Brain: A Journal of Neurology, 133*(1), 234–247. https://doi.org/10.1093/brain/awp269

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science, 27*(3), 525–559. https://doi.org/10.1016/S0364-0213(03)00011-9

Schulz, D., Richter, T., Schindler, J., Lenhard, W., & Mangold, M. (2023). Using accuracy and response times to assess inhibitory control in kindergarten children: An analysis with explanatory item response models. *Journal of Cognition and Development, 24*(1), 82–104. https://doi.org/10.1080/15248372.2022.2119977

Simpson, A., & Carroll, D. J. (2019). Understanding early inhibitory development: Distinguishing two ways that children use inhibitory control. *Child Development, 90*(5), 1459-1473. https://doi.org/10.1111/cdev.13283

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience, 9*, Article 102. https://doi.org/10.1186/1471-2202-9-102

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(10), 1231–1238. https://doi.org/10.1097/00004583-200210000-00013

Tellegen, P. J., Laros, J. A., & Petermann, F. (2006). *Deutsche Normierung und Validierung des SON-R 2½ - 7* [Norming and validation of the SON-R 2½ – 7 in German]. Hogrefe.

Thorell, L. B., Bohlin, G., & Rydell, A.-M. (2004). Two types of inhibitory control: Predictive relations to social functioning. *International Journal of Behavioral Development, 28*(3), 193–203. https://doi.org/10.1080/01650250344000389

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58*(1), 140–159. https://doi.org/10.1016/j.jml.2007.04.006

Yücel, M., Fornito, A., Youssef, G., Dwyer, D., Whittle, S., Wood, S. J., Lubman, D. I., Simmons, J., Pantelis, C., & Allen, N. B. (2012). Inhibitory control in young adolescents: The role of sex, intelligence, and temperament. *Neuropsychology, 26*(3), 347-356. https://doi.org/10.1037/a0027693

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., & 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

# Chapter VI

# General Discussion

## Summary and Key Findings Across the Three Studies

The findings from the three studies conducted in context of this dissertation shed light on the complex nature of inhibitory control development in children and provide insights into the challenges associated with its assessment. Significantly, each study produced actionable recommendations. In the first study, I investigated the validity of response accuracy and latency within the computerized Pointing Stroop Task and explored whether a combined analysis of these metrics offers additional value compared to examining them separately. The results supported the idea that the used instrument validly measures inhibitory control in terms of response accuracy and latency and underlined the importance of considering both measures in assessing inhibitory control. The significant positive association between inhibitory control and fluid intelligence and the fact that the interaction of both metrics delivers an explanation of variance above the main effect of the two metrics further supports the validity of the integrative approach towards response accuracy and latency.

Building upon these findings, the second study investigated the relationship between response accuracy, response latency, and inhibitory control in a cross-sectional analysis using the validated test. This study revealed significant differences in the relationship between response accuracy and latency across different developmental stages. Younger children showed a more pronounced benefit in terms of response accuracy when taking more time to answer. This result suggests that they might need more time to inhibit their prepotent (and in the context of inhibitory control wrong) responses. In contrast, older children exhibited a weaker association, indicating they had developed more efficient inhibitory mechanisms. In fact, the results suggest that for younger children, who still profit from taking more time, longer response latencies might be desirable, as they indicate the ability to inhibit the prepotent response long enough to answer correctly. These results highlight the non-linear development of inhibitory control and the importance of considering age-related differences when assessing inhibitory control in children. From a practical standpoint, I discovered that specific metrics demonstrated more robust correlations with recognized covariates, leading to clear metric recommendations. Additionally, the findings suggest that if researchers aim to employ response latency as a marker of inhibitory control in very young children, they should prioritize latencies from correct responses. This supports a prevalent methodology in the domain of inhibitory control development (e.g., Camerota et al., 2020; Davidson et al., 2006; Magnus et al., 2019; Ursache & Raver, 2014); a method which, to my understanding, has not been previously examined empirically.

The third study focused on the longitudinal development of inhibitory control and the relationship between response accuracy and latency over time. This study revealed a dynamic relationship between response accuracy, latency, and inhibitory control. Children, who were able to answer accurately early, improved their response latency faster. That is, their responses got faster quicker, than for children who still needed time to develop the ability to answer accurately. This indicates that children start developing their response latency more efficiently when they already have achieved the ability to answer accurately. Children who attained high accuracy at younger ages initially took more time to respond, indicating that this intentional pacing is beneficial. This deliberate early approach meets the central aim of inhibitory control tasks – accuracy – and sets the stage for them to develop faster response latencies while retaining accuracy as they age. Beyond these developmental results, the study also found that response latency was only negatively related to fluid intelligence if accuracy remained high. This finding further indicates that speeding up and answering quickly is only desirable for children who have developed the ability to answer correctly. These observations underscore a tangible recommendation for practice: prioritize improving or training accuracy during the initial stages and shift emphasis to response latency and speed only once consistent accuracy has been established and sustained.

Overall, the findings from the three studies highlight the complex and non-linear nature of inhibitory control development in children. They emphasize the importance of considering both response accuracy and latency in assessing inhibitory control and provide insights into the developmental trajectories and age-related differences in inhibitory control. The dynamic relationship between response accuracy, latency, and inhibitory control over time underscores the need for longitudinal assessments to capture the changes in inhibitory control abilities as children develop. Furthermore, the tangible recommendations drawn from the studies highlight the value of this research for both academic and practical domains. Emphasizing accuracy in the early stages and gradually shifting focus to speed as children mature provides a structured approach for educators, clinicians, and researchers working to optimize inhibitory control training and assessment in children.

## Integration into and Extension of Current Literature

In this section, I will integrate the findings of our empirical studies with the existing literature on inhibitory control and its measurement. I will examine how the results align with previous research and contribute to and extend the current understanding of inhibitory control in children.

### Response Accuracy and Latency as Indicators of Inhibitory Control Ability

One of the primary goals of our dissertation was to assess the validity of using response accuracy and latency as indicators of inhibitory control in young children. Our results support the reliability of these metrics as measures of inhibitory control. In line with the literature (e.g., Paap, 2019), the data suggests that response accuracy is a preferred indicator as long as it effectively discriminates between participants. Only when response accuracy no longer has enough discriminatory value – for instance, when children reach a performance ceiling (Roebers, 2017) or achieve sufficiently high accuracy (Zelazo et al., 2013) – should response latency be incorporated. This approach builds upon prior research that either primarily focuses on a single metric (Petersen et al., 2016) or underscores the significance of integrating both accuracy and latency when gauging inhibitory control (Camerota et al., 2019; Camerota et al., 2020; Magnus et al., 2019; Zelazo et al., 2013). Moreover, the direct juxtaposition of different methods – whether model-based (Magnus et al., 2019), formula-based (Liesefeld & Janczyk, 2019; Vandierendonck, 2017) or employing straightforward raw scores – facilitated comprehensive comparisons, resulting in specific recommendations. Reinforcing this position, the longitudinal structure of the third study affirms the notion that integrating response latency with accuracy must be approached with an understanding of the non-linear trajectory of inhibitory control development (Camerota et al., 2020). This pattern of results demonstrates that the importance of faster response times becomes apparent and relevant only after consistently achieving high accuracy.

**Inhibitory Control and Intelligence**

Consistent with prior research (e.g., Yücel et al., 2012) and theory (Dempster, 1991), we found a association between inhibitory control and fluid intelligence. The association suggests that inhibitory control, measured by the computerized Pointing Stroop Task, is linked with higher-order cognitive processes such as reasoning, problem-solving, and abstract thinking (Horn & Cattell, 1967). Notably, the association was most pronounced when both metrics – response accuracy and latency – were integrated. The literature has not fully elaborated the relationship between fluid intelligence and inhibitory control. Some studies indicate that inhibitory control has predictive value for fluid intelligence but may not offer additional value beyond other executive functions (Duan et al., 2010; Friedman et al., 2006). Yet, when the combined approach of response accuracy and latency is applied in young children, as outlined in this dissertation, the specificity of the inhibitory control association could become more distinct; a question that might be worth further exploration in subsequent studies.

**Challenges in Assessing Inhibitory Control: Rapid and Non-linear Development**

Another core question was to investigate the intricacies of assessing inhibitory control, mainly focusing on its swift and non-linear development. The aim was to decipher the differential manifestation of inhibitory control across developmental stages, skill sets, and the evolving relationship among response accuracy, latency, and inhibitory control as children mature.

Aligned with existing literature, we observed that inhibitory control in young children undergoes a swift (Kochanska et al., 1996) and non-linear trajectory (Camerota et al., 2020). This observation underscores the necessity of factoring in developmental stages and individual variances when assessing inhibitory control. Younger participants showed elongated response latencies relative to their older counterparts. Yet, for these younger children, prolonged response latencies might not necessarily indicate poor inhibitory control; instead, accuracy serves as a more definitive metric. Our analysis further illuminated the dynamic relationship among response latency, accuracy, and inhibitory control. Evaluating the entirety of response latencies, including those not linked to correct answers, unveiled that younger children with slower latencies often exhibited enhanced response accuracy. This finding suggests an intentional, contemplative approach to tasks necessitating inhibitory control, very much in line with the definition of inhibitory control (Verbruggen & Logan, 2008). Conversely, older children showcased swifter latencies and heightened accuracy, indicative of a more streamlined and ingrained inhibitory control mechanism. These observations resonate with the horse race model (Logan & Cowan, 1984), proposing that as children grow and develop, they gradually acquire the ability to swiftly invoke and sustain inhibitory processes, thereby effectively suppressing dominant responses. Further enriching this discourse, applying the dual processing theory (Schneider & Fisk, 1983) within the realm of young children’s inhibitory control offered granular insights into its cognitive mechanics.

Interestingly, the narrative shifts when exclusively considering response latencies from accurate answers. When considering only response latencies from correct responses, the association between response latency and inhibitory control ability is more consistent across the full age range. Therefore, such latencies seemingly provide a more precise portrayal of inhibitory control across the age span of young children. When children falter in their responses, it remains unclear whether inhibitory processes were initiated. If initiated but unsuccessful, it suggests the inhibitory action was either initiated too late (Logan & Cowan, 1984) or lacked sufficient intensity or duration (Simpson & Carroll, 2019). Yet, in tasks like the computerized Pointing Stroop Task, where children have a 50% probability of guessing correctly, a child who never attempts to inhibit would still register accurate answers in half the instances. This fact complicates the interpretation, especially when employing model-based methodologies to gauge inhibitory control (e.g., Camerota et al., 2020; Magnus et al., 2019) – missing data patterns, in this context, present analytical challenges. While our study reinforced existing methodologies and literature, it also contributed novel, pragmatic recommendations.

### *Missing Data in Inhibitory Control Assessments: Analytical Considerations*

Research in inhibitory control in young children heavily depends on the quality and nature of missing data. Notably, the absence of data in response latencies of incorrect responses may influence our understanding of inhibitory control in children. In our studies, it has been shown that very young children especially make a non-negligible number of errors. If one excludes all response latencies from that trial, a substantial (more than 30% in our study) pattern (dependent on age) emerges.

Data missing completely at random (MCAR) implies a non-systematic pattern. However, if missing at random (MAR), the absence is linked to observable variables. Most critically, when data are missing not at random (MNAR), they're correlated with unseen values, posing analysis hurdles (Rubin, 1976). Our observation of more prevalent missing data in younger participants possibly deviates from MCAR, suggesting age-related systematic variations that need exploration. Structural equation models using the full information maximum likelihood estimator (FIML) can manage significant missing data, but their success depends on the pattern of missingness (Enders, 2023). Although FIML can handle up to 20-30% of MAR missing data efficiently (Graham, 2009), the distribution matters. As with our younger participants, a clustered absence in a subgroup may reduce FIML's effectiveness (Collins et al., 2001; Schafer & Graham, 2002).

Not explicitly considering age as a factor within one's dataset could introduce potential methodological pitfalls. When missing data correlates primarily with skill, and item-level response accuracies are incorporated into the analysis, the dataset could be declared MAR, the premise being that all foundational data dictating the missingness of response latency is embedded within the model. However, age possibly wields influence beyond mere response accuracy metrics. The interplay of other executive functions, known to share variance with inhibitory control (Friedman et al., 2006), combined with developmental trajectories, could further dictate data missingness patterns. For instance, circumstances necessitating the exclusion of response latencies – due to excessively swift reactions (e.g., Mangus et al., 2019) or prolonged latencies (our studies) – as well as instances where items had to be omitted because of participant distraction, lack of initial task comprehension, or external influences (e.g., assistance from research assistants or parents, as evidenced in some of our studies), would introduce missing data not exclusively determined by response accuracy and possibly dependent on age. Should such factors substantially shape the absent data landscape, the argument for MNAR gains traction. The limitations of FIML in effectively addressing MNAR are well-documented (Enders & Bandalos, 2001). Our study supports prevailing practices, suggesting that response latencies from correct answers offer a more precise lens into inhibitory control. Nonetheless, a more nuanced exploration is warranted to fully grasp the repercussions of omitting response latencies associated with incorrect answers.

### *Meaning of Response Latencies from Incorrect Responses*

The present work corroborated that quicker response latencies for correct answers are more indicative of inhibitory control than latencies for incorrect answers. Yet, this observation prompts a pivotal question for further research: Do response latencies contain valuable information for assessment or diagnostic purposes, even if they are associated with incorrect responses? The time-on-task analysis from this research suggests that the ability to delay a response, even one that is ultimately incorrect, might reflect the initiation and maintenance of an inhibitory process. As per Logan and Cowan's (1984) theory, even if this process does not result in a correct response, it may still represent the initiation of an inhibition. This hypothesis, generated by our current findings, provides a novel contribution to the field, indicating that the role of incorrect response latencies deserves closer examination. Future studies are encouraged to investigate this potential, which could lead to a more nuanced understanding of inhibitory control across developmental stages.

## Development of Inhibitory Control: Longitudinal Development

The longitudinal perspective adopted in the third study of this dissertation delivers important insight into the developmental patterns of inhibitory control among children, adding a valuable dimension to our knowledge that cannot be thoroughly investigated by cross-sectional studies alone. Through capturing response accuracies and latencies over time, we obtain a dynamic picture of inhibitory control, moving from a domain largely governed by accuracy in early stages to one in which latency gains importance as children mature.

The findings from our studies align closely with the developmental trajectories proposed by previous research, which suggests a rapid (Kochanska et al., 1996) non-linear

(Camerota et al., 2020) maturation of executive functions, including inhibitory control. Our research complements these findings by considering the nuanced interplay between accuracy and latency in the development of inhibitory control and the implications for children's cognitive assessments. Notably, the observed pattern – where children who achieve accuracy early on show an faster improvement in response latency – underscores the importance of considering both accuracy and latency in longitudinal assessments. Children who have not yet mastered accurate responses appear to benefit from taking their time using longer latencies for controlled inhibition. As they achieve consistently accurate responses, they are better positioned to focus on improving response latency without sacrificing accuracy.

This developmental shift in the importance of latency is consistent with the theoretical framework of cognitive efficiency, which posits that as children's cognitive processes become more automatic, they can perform tasks faster without compromising performance, another aspect indicating similarities with intelligence (cf. Neubauer & Fink, 2009). It also relates to the speed-accuracy tradeoff (Heitz, 2014) described in the literature, initially suggesting that younger children should tend to prioritize accuracy at the expense of speed, only to switch their focus as their executive functions consolidate.

Overall, this dissertation significantly advances our understanding of inhibitory control by highlighting the longitudinal dynamics of response accuracy and latency. It reveals how the interplay between these metrics evolves throughout childhood, with accuracy being pivotal in the early stages and latency gaining importance as children grow and develop their skills. This research offers a nuanced perspective, integrating empirical evidence cross-sectionally and longitudinally to formulate actionable insights for research. It underlines the necessity of a tailored approach in evaluating inhibitory control in children.

## Practical Implications

The empirical findings uncovered in this dissertation have implications for both the clinical and educational setting, especially when creating programs and interventions intended to foster inhibitory control in young children. Understanding the developmental complexities and subtleties in measuring inhibitory control can greatly enhance the design and efficacy of these initiatives.

### Clinical Setting

The validation of both response accuracy and latency as integral components of inhibitory control measures offers a more robust framework for the evaluation of a concept with relevance in both clinical (e.g., Iacono et al., 2008) and educational (e.g., Raver & Blair, 2016) settings. Clinicians may be able to use our findings to better assess and monitor the development of inhibitory control in children. The recommendation to only look at speed after a certain accuracy was reached and can be maintained and our explicit recommendation for using some assessment methods can help guide the decision process of clinicians, both in practice and research. In particular, we recommended only considering response latency after children have reached an accuracy threshold, or, if they want to include response latency for children not reaching this threshold, modeling inhibitory control as a combination of response accuracy and latency only, including response latencies from correct response. While the latter procedure is standard (e.g., Camerota et al., 2020; Davidson et al., 2006; Magnus et al., 2019; Ursache & Raver, 2014), it has, to our knowledge, never been tested and our work is the first to underpin this reasoning empirically.

### Educational Setting

In the context of education, the findings from the present work suggest that educators can design activities and learning material that emphasize both response accuracy and latency. For instance, educators can create games that require children to inhibit their immediate responses and take time to respond accurately. The challenge level of these games can be tailored to age and developmental stage, gradually increasing over time and shifting from aiming for accurate (effective) to speed-related (efficient) aims as children improve their inhibitory control abilities. For example, suppose it becomes evident that the child does not yet have the ability to maintain high accuracy when starting to speed up. In that case, interventions tailoring effective inhibition should remain the main focus.

**Interventions**

Both the clinical and the educational areas include interventions. Our findings underscore the importance of early interventions. If children are given opportunities to develop inhibitory control skills in terms of accuracy at an early age, they are likely to develop quicker, possibly helping them in behavioral areas (e.g., Rhoades et al., 2009) and with cognitive capabilities (e.g., Wilkinson et al., 2020). Early interventions may also help children who are struggling with inhibitory control, helping them improve their abilities over time. As shown in study 3, if children can inhibit long enough to answer correctly at an early age, their efficiency can also develop more quickly, even though they might be slower at the start. Educators might keep this result in mind when supervising the developmental process of children in terms of inhibitory control.

Recent training approaches have effectively developed inhibitory control, with evidence of transfer to real-world situations. For example, Wilkinson et al. (2020) demonstrated the beneficial role of inhibitory control training in mathematics and science education. By embedding inhibitory control exercises within the content of these subjects, a Stop & Think intervention was applied, in which children played a gameshow in which they were to stop for a short moment before answering. Beyond that, the possibility to answer was locked for 5 seconds during this training gameshow. Wilkinson et al. (2020) showed that the intervention partially improved children's counterintuitive reasoning and academic development. Nonetheless, the findings of the present work show an opportunity to refine such interventions further, potentially improving their efficacy through a more nuanced application of response latency considerations.

The dissertation's results highlight a developmental arc where younger children benefit from longer response latencies, allowing more deliberate, accurate responses. As children mature, the role of response latency becomes more pronounced, and a capacity for quicker responses signals the consolidation of inhibitory control mechanisms. This maturation of response speed, duly calibrated with maintained response accuracy, indicates the development of efficient inhibitory control, which, according to our findings, is closely associated with higher cognitive functions such as fluid intelligence.

Considering these findings, implementing adaptive interventions is proposed for further study and potential practical application. An adaptive intervention that permits younger or less proficient children more time to answer, thereby strengthening their early inhibitory processing with an emphasis on accuracy, would emphasize accuracy over latency at the early stages of inhibitory control development. As children progress, dynamic pacing for response latencies could be introduced, where the time allowed for responses gradually decreases, scaffolding their capacity for quicker, yet still accurate, reactions. Furthermore, adaptive software monitoring could support individualized learning trajectories (e.g., Christodoulou & Angeli, 2022) and adjusting the balance between accuracy and speed, tailoring challenges to the child's current proficiency level.

By incorporating these suggested enhancements, interventions like Stop & Think can be more closely aligned with developmental science, providing a robust scaffold for children's cognitive development, and amplifying the impact of inhibitory control training within educational settings. Integrating the complexity revealed by the research into practical applications promises a more nuanced approach to fostering the intellectual growth of children and optimizing their success in mastering counterintuitive concepts and a broader spectrum of academic challenges.

More generally, by considering these two parameters, educators, clinicians, and researchers can better understand a child's inhibitory control capabilities. The dual utilization of these measures is vital for identifying children who may require targeted interventions to improve their inhibitory control skills, thus potentially mitigating academic (e.g., Ng et al., 2014) and social difficulties (e.g., Rhoades et al., 2009) early on.

## Implications for Future Research and Methodological Reflections

In the following section, I will go into the implications arising from this dissertation for research in inhibitory control among young children, broader cognitive capabilities, and methodological approaches in the field.

### Integrating Binary Accuracy and Continuous Latency

We encounter a methodological challenge when evaluating inhibitory control due to the inherent differences between response accuracy and response latency. Response accuracy offers a binary outcome, categorized as either "correct" or "incorrect". It provides a precise, straightforward measure of an individual's performance but lacks depth regarding the nuances of how the task was performed. Conversely, response latency is a continuous measure, capturing the time taken to respond. This difference introduces complexity, particularly when correlating or integrating it with the binary data from response accuracy.

The challenge, then, is to derive meaningful insights from these two distinct data types. On one hand, we have the clear-cut, binary nature of accuracy; on the other, the more nuanced, continuous measure of latency. Combined, they can offer a holistic view of inhibitory control. Still, the methodological challenge lies in ensuring that this integration is both meaningful and representative of the underlying cognitive processes.

The challenge intensifies when considering the inherent speed-accuracy tradeoff (Heitz, 2014). In many cognitive tasks, particularly those assessing inhibitory control, individuals might prioritize speed over accuracy, "rushing through" the task to finish quickly rather than accurately. Conversely, others might adopt a more conservative strategy, ensuring they respond correctly even if it means taking more time. An individual with fast response latencies but low accuracy might be inaccurately perceived as having strong inhibitory control when, in reality, they're just rushing through the task. Attempting to integrate the binary accuracy data with the continuous latency data without considering the speed-accuracy tradeoff can lead to misleading conclusions.

In addressing the intricate challenge of integrating binary accuracy and continuous latency to assess inhibitory control in young children, we have added a nuanced understanding of how the developmental trajectory of inhibitory control impacts the interpretability of response latency as a diagnostic measure. By empirically demonstrating the non-linear relationship between response accuracy, latency, and inhibitory control ability, our work highlights the complexities inherent in measuring these cognitive processes across a broad age spectrum. We have also underscored the criticality of achieving a certain threshold of accuracy before response latency can serve as a reliable proxy for inhibitory control proficiency. To this end, we applied model-based approaches, allowing for deeper insights into the distinct contributions of item-level response accuracies and latencies. By employing careful analyses, including time-on-task effects and considering age as a moderating factor, we were able to differentiate the conditions under which response accuracies and latencies provide meaningful information about inhibitory control.

The findings of the present dissertation can be instrumental for researchers and practitioners aiming to assess and support the development of inhibitory control in children. We provide a framework for interpreting response latencies in relation to response accuracy, emphasizing the importance of establishing an accuracy threshold before considering response speed as indicative of inhibitory control competence.

**Reflection of Applied Methodological Approaches**

Throughout our studies, various approaches have been employed, from composite measures (Liesefeld & Janczyk, 2019; Vandierendonck, 2017) that combine multiple indicators into a single score to model-based approaches (Camerota et al., 2020; Magnus et al., 2019) that allow a more nuanced investigation. Each approach has its strengths. For instance, composite measures provide a straightforward, easily applicable, and sample-independent way of integrating both metrics. Conversely, model-based approaches offer deeper insights, providing specific interpretations of item-level response accuracies and response latencies. However, the limitations of these methods cannot be ignored. Composite measures may oversimplify, potentially overlooking subtleties, while model-based approaches are complex and have only been tested on single samples. As estimates are made for the specific sample, it has not yet been validated on a separate sample. Therefore, the value of model-based approaches has so far been in more deeply investigating inhibitory control as in the second study of the present work and current research (Camerota et al., 2020; Magnus et al., 2019), not assessing it, for example, in a clinical setting.

**Implications for Using Inhibitory Control Tests Across a Broader Age Range**

One pivotal aspect of our investigation into inhibitory control was exploring the validity of employing a single test across a broad age range. When only considering response accuracy, the validity of such a test can be limited to a very narrow age range (Petersen et al., 2016). This view posits that a given test may be too complex for younger children while simultaneously failing to challenge older children sufficiently, potentially leading to a ceiling effect where older children perform at their highest level, making it difficult to distinguish subtle differences in capability (Roebers, 2017). While our results support that a ceiling effect is quickly reached for the older subsample, the present work opens up the potential to stretch the age-appropriateness of such tasks, offering a deeper insight into inhibitory control development. Specifically, our results indicate that integrating response accuracy and latency into the assessment process significantly enhances inhibitory control measures' sensitivity and discriminative power. This result implies that assessments can be adapted and applied effectively over a more extended developmental period than previously assumed.

Our approach capitalizes on the non-linear trajectory of inhibitory control development, with younger children benefiting from tasks that assess accuracy in conjunction with latency, fostering a deliberate and controlled processing approach. As children mature, the focus transitions from solely accuracy-based measures to those that simultaneously account for the efficiency of response – a reflection of the increased proficiency in inhibitory processes. This dynamic assessment strategy, informed by our empirical findings, gives credence to the potential of using a single test, such as the computerized Pointing Stroop Task (Berger et al., 2000), as a valuable and valid tool for a more extended developmental period. By employing a measurement model that incorporates both binary response accuracy and continuous response latency, tailoring to the children's specific developmental stages, our research lays the groundwork for extending the utility of single-task assessments.

In essence, the revised analytical lens guided by our research allows for an adaptable assessment metric sensitive to the evolving proficiency levels in inhibitory control across different ages. However, it is necessary to caution that a nuanced and individualized approach is still warranted despite these advancements. While a single test can apply over an extended age range, it is vital to consider developmental appropriateness on a case-by-case basis, ensuring tasks remain challenging and informative. Nonetheless, our findings are a promising step toward more valid, efficient, and robust assessments of inhibitory control that can withstand the vast heterogeneity encountered in developmental psychology.

**Future Research Directions**

Building upon the findings of this dissertation, multiple areas for future research become obvious, offering the potential to enhance our understanding and practical application of inhibitory control in early childhood.

Following the longitudinal perspectives established, a subsequent investigation could look deeper into the developmental trajectory of inhibitory control. Given the nuanced relationship between response latency and accuracy observed, future studies might scrutinize this interaction in varied contexts, such as socio-economically diverse settings or in populations with specific educational needs. How does inhibitory control development differ among children with varying access to resources, learning support, or language background? Understanding these dimensions could shed light on customizing educational interventions more effectively. Additionally, empirical work could focus on the influence of inhibitory control on educational outcomes. Rigorous interventions that aim to bolster inhibitory control could be assessed longitudinally for their impact on academic performance and adaptability in learning environments. Such research might consider individualized instructional methodologies, considering cognitive and environmental factors that could moderate these effects.

The insights from this dissertation suggest that task complexity could significantly impact children's response patterns. Hence, future research could systematically vary the complexity of inhibitory control tasks to dissect how this factor influences the speed-accuracy tradeoff in different age groups. This approach would involve manipulating both the cognitive load and the nature of distractions presented during tasks that measure inhibitory control.

Building on our current understanding, additional research might target specific mechanisms within inhibitory control (c.f., Harnishfeger, 1995; Nigg, 2000), across developmental stages. Employing a range of tasks known to isolate these mechanisms, researchers could compare the maturation of these facets in early childhood, potentially leading to more targeted assessment and intervention techniques.

Lastly, to translate research findings into practical strategies, future studies should aim to improve the application of inhibitory control assessments in classroom and clinical settings. This next step could involve developing and validating tools that are theoretically sound and practically feasible for educators and clinicians to implement.

This dissertation lays the groundwork for these diverse directions and underscores the importance of multidisciplinary research endeavors. By investigating these directions, further research can enrich our understanding of inhibitory control development. It may lead to applicable interventions that could foster the cognitive growth and well-being of children.

## Dissertation Limitations

While this dissertation has yielded significant insights into the development of inhibitory control in children, it is essential to consider its limitations to contextualize the conclusions drawn.

The studies reported in this dissertation did not include assessments of other executive functions. Executive functions encompass a variety of cognitive processes necessary for goal-directed behavior, including working memory, cognitive flexibility, and inhibitory control. Since these functions are closely intertwined and often work simultaneously (Friedman et al., 2006), focusing only on inhibitory control may have provided a somewhat isolated view of children's cognitive capacities. However, the identification and detailed examination of inhibitory control was the dissertation's primary aim. Future research would benefit from considering other executive functions in studying children's cognitive development to provide a more comprehensive picture. Further, the dissertation relied solely on the Pointing Stroop Task (Berger et al., 2000) in measuring inhibitory control. While this task is well-validated and its validity reinforced within this dissertation, exclusive reliance on a single instrument may constitute a limitation.

The sample employed in these studies was relatively heterogeneous regarding age, language, and cultural backgrounds, taken from two countries (Switzerland and Germany). However, it is important to note that the sample predominantly comprised of participants from high socio-economic status (SES) backgrounds. Due to the association between SES and the two main measures used in this study, namely inhibitory control (Kałamała et al., 2020; Lipina et al., 2013) and fluid intelligence (though less than with crystallized intelligence; Rindermann et al., 2010), we may have gotten unrepresentative high ability responses, because our SES was high for most children. As such, the findings may not generalize to children from lower SES backgrounds. Future studies should aim to include a more socio-economically diverse sample to broaden the applicability of the results.

It is also noteworthy that the testing period spanned from October 2019 to April 2023. During that time, in Germany (Ravens-Sieberer et al., 2022) and Switzerland (Bringolf-Isler et al., 2021), strict measures were taken to combat the COVID-19 pandemic. For the present studies presented in this dissertation, the result was that data collection was to be carried out under diverse circumstances, like being separated from the experimenter by a plastic glass or wearing a mask. Further, one child could have been tested at one measurement point with the restrictions and at another without these restrictions. Another potential confounding fact is that the sample was recruited in both Germany and Switzerland, which had very different restrictions. While all protocols were adhered to, the pandemic's unavoidable influence potentially added an additional layer of variance in the data. Testing this variance, however, would be out of the scope of this dissertation.

Lastly, although this dissertation has provided a detailed examination of the development of inhibitory control, it focused primarily on cold response inhibition. Given that inhibitory control is a multi-faceted construct encompassing a variety of capacities (Harnishfeger, 1995; Nigg, 2000), this limited focus may not have captured the full complexity of inhibitory control development. While this approach allows one to investigate very specifically and learn precise knowledge, it might not be very generalizable. Future research could probe these different facets of inhibitory control for a more nuanced understanding of this complex ability. For instance, future studies might explore whether the response accuracy-latency pattern observed in this research also manifests in other aspects of inhibitory control, such as hot response inhibition (see Simpson & Carroll, 2019), when assessed using a delay of gratification task.

## Concluding Remarks and Future Directions

The work presented in this dissertation has advanced our understanding of inhibitory control's developmental nuances in early childhood. Through the three studies that span both cross-sectional and longitudinal analyses, this dissertation highlighted the evolving interplay between response accuracy and latency in assessing and developing inhibitory control. By validating these metrics and exploring their relationship with cognitive development, the findings equip researchers, clinicians, and educators with a more comprehensive understanding of this essential executive function. This exploration of inhibitory control's rapid and non-linear development underscores the importance of tailored approaches in assessment and intervention. The methodological advancements and empirical insights should encourage future research designed to fine-tune the measurement and deeper understanding of the development of inhibitory control.

In conclusion, this dissertation primarily enhanced the academic comprehension of inhibitory control, but it also could influence practical applications, either directly or through subsequent research. By emphasizing developmentally appropriate assessments and early interventions, this research may foster improved educational practices and clinical interventions to better support young children's cognitive development.

**References**

Berger, A., Jones, L., Rothbart, M. K., & Posner, M. I. (2000). Computerized games to study the development of attention in childhood. *Behavior Research Methods, Instruments, & Computers, 32*(2), 297-303. https://doi.org/10.3758/BF03207798

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin, 143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bialystok, E. & Martin, M.M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325–339.

Bringolf-Isler, B., Hänggi, J., Kayser, B., Suggs, L. S., Dössegger, A., Probst-Hensch, N., & SOPHYA-Study Group (2021). COVID-19 pandemic and health related quality of life in primary school children in Switzerland: A repeated cross-sectional study. *Swiss Medical Weekly, 151*, w30071. https://doi.org/10.4414/smw.2021.w30071

Camerota, M., Willoughby, M. T., & Blair, C. B. (2019). Speed and accuracy on the Hearts and Flowers task interact to predict child outcomes. *Psychological Assessment, 31*(8), 995–1005. https://doi.org/10.1037/pas0000725

Camerota, M., Willoughby, M. T., Magnus, B. E., & Blair, C. B. (2020). Leveraging item accuracy and reaction time to improve measurement of child executive function ability. *Psychological Assessment, 32*(12), 1118–1132. https://doi.org/10.1037/pas0000953

Christodoulou, A., & Angeli, C. (2022). Adaptive learning techniques for a personalized educational software in developing teachers' technological pedagogical content knowledge. *Frontiers in Education, 7*, 789397. https://doi.org/10.3389/feduc.2022.789397

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias?. *Psychological Science, 26*(1), 99–107. https://doi.org/10.1177/0956797614557866

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*(2), 157-173. https://doi.org/10.1016/0160-2896(91)90028-C

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year-old children. *Psychological Test and Assessment Modeling, 52*(4), 419–431.

Enders, C. K. (2023). Fitting structural equation modeling with missing data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed., pp. 223-240). Guilford.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). Academic Press. https://doi.org/10.1016/B978-012208930-5/50007-6

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*, 150. https://doi.org/10.3389/fnins.2014.00150

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*(2), 107–129. https://doi.org/10.1016/0001-6918(67)90011-X

Iacono, W. G., Malone, S. M., & McGue, M. (2008). Behavioral disinhibition and the development of early-onset addiction: Common and specific influences. *Annual Review of Clinical Psychology, 4*, 325–348. https://doi.org/10.1146/annurev.clinpsy.4.022007.141157

Johann, V. E., Enke, S., Gunzenhauser, C., Könen, T., Saalbach, H., & Karbach, J. (2022). Executive functions in mono- and bilingual children: Factor structure and relations with fluid intelligence. *Journal of Experimental Child Psychology, 224*, 105515. https://doi.org/10.1016/j.jecp.2022.105515

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition, 204*, 104373. https://doi.org/10.1016/j.cognition.2020.104373

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490-507. https://doi.org/10.2307/1131828

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Lipina, S., Segretin, S., Hermida, J., Prats, L., Fracchia, C., Camelo, J. L., & Colombo, J. (2013). Linking childhood poverty and cognition: Environmental mediators of non-verbal executive control in an Argentine sample. *Developmental Science, 16*(5), 697–707. https://doi.org/10.1111/desc.12080

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Magnus, B. E., Willoughby, M. T., Blair, C. B., & Kuhn, L. J. (2019). Integrating item accuracy and reaction time to improve the measurement of inhibitory control abilities in early childhood. *Assessment, 26*(7), 1296–1306. https://doi.org/10.1177/1073191117740953

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience and Biobehavioral Reviews, 33*(7), 1004–1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

Ng, F. F.-Y., Tamis-LeMonda, C. S., Yoshikawa, H., & Sze, I. N.-L. (2014). Inhibitory control in preschool predicts early math skills in first grade: Evidence from an ethnically diverse sample. *International Journal of Behavioral Development, 39*(2), 139–149. https://doi.org/10.1177/0165025414538558

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Paap, K. R. (2019). The bilingual advantage debate: Quantity and quality of the evidence. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 701–735). Wiley-Blackwell. https://doi.org/10.1002/9781119387725.ch34

Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., & Staples, A. D. (2016). Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review, 40*, 25–71. https://doi.org/10.1016/j.dr.2016.02.001

Ravens-Sieberer, U., Kaman, A., Erhart, M., Devine, J., Schlack, R., & Otto, C. (2022). Impact of the COVID-19 pandemic on quality of life and mental health in children and adolescents in Germany. *European Child & Adolescent Psychiatry, 31*(6), 879-889. https://doi.org/10.1007/s00787-021-01726-5

Raver, C. C., & Blair, C. (2016). Neuroscientific insights: Attention, working memory, and inhibitory control. *The Future of Children, 26*(2), 95–118. https://doi.org/10.1353/foc.2016.0014

Rhoades, B. L., Greenberg, M. T., & Domitrovich, C. E. (2009). The contribution of inhibitory control to preschoolers' social–emotional competence. *Journal of Applied Developmental Psychology, 30*(3), 310–320. https://doi.org/10.1016/j.appdev.2008.12.012

Rindermann, H., Flores-Mendoza, C., & Mansur-Alves, M. (2010). Reciprocal effects between fluid and crystallized intelligence and their dependence on parents' socioeconomic status and education. *Learning and Individual Differences, 20*(5), 544–548. https://doi.org/10.1016/j.lindif.2010.07.002

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581-592. https://doi.org/10.1093/biomet/63.3.581

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Simpson, A., & Carroll, D. J. (2019). Understanding early inhibitory development: Distinguishing two ways that children use inhibitory control. *Child Development, 90*(5), 1459-1473. https://doi.org/10.1111/cdev.13283

Schneider, W., & Fisk, A. D. (1983). Attentional theory and mechanisms for skilled performance. In R. A. Magill (Ed.), *Memory and control of action* (pp. 119–143). North-Holland. https://doi.org/10.1016/ S0166-4115(08)61989-5

Ursache, A., & Raver, C. C. (2014). Trait and state anxiety: Relations to executive functioning in an at-risk sample. *Cognition and Emotion, 28*(5), 845–855. https://doi.org/10.1080/02699931.2013.855173

van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S. H., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences, 9*(3), 27. https://doi.org/10.3390/bs9030027

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Wilkinson, H. R., Smid, C., Morris, S., Farran, E. K., Dumontheil, I., Mayer, S., Tolmie, A., Bell, D., Porayska-Pomsta, K., Holmes, W., Mareschal, D., Thomas, M. S. C., & UnLocke Team (2020). Domain-specific inhibitory control training to improve children's learning of counterintuitive concepts in mathematics and science. *Journal of Cognitive Enhancement: Towards the Integration of Theory and Practice, 4*(3), 296–314. https://doi.org/10.1007/s41465-019-00161-4

Yücel, M., Fornito, A., Youssef, G., Dwyer, D., Whittle, S., Wood, S. J., Lubman, D. I., Simmons, J., Pantelis, C., & Allen, N. B. (2012). Inhibitory control in young adolescents: The role of sex, intelligence, and temperament. *Neuropsychology, 26*(3), 347-356. https://doi.org/10.1037/a0027693

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

**Acknowledgements**

Diese Arbeit wäre ohne die Unterstützung und Mitwirkung vieler Menschen nicht zustande gekommen, und es ist mir eine große Freude, ihnen an dieser Stelle meinen Dank auszusprechen.

Mein aufrichtiger Dank gilt Herrn Prof. Dr. Tobias Richter, dessen umfassende Unterstützung, wertvolle Rückmeldungen und fachliche Expertise entscheidend zum Erfolg dieser Arbeit beigetragen haben. Ich bin auch meinen Koautorinnen und Koautoren – Julia Schindler, Wolfgang Lenhard und Robin Segerer – zu großem Dank verpflichtet. Besonderer Dank gilt Madlen Mangold, deren fachliche und menschliche Unterstützung maßgeblich für den Erfolg dieser Arbeit war. Außerdem möchte ich mich bei PD Dr. Simon Tiffin-Richards und PD Dr. Eva Michel für die Begutachtung dieser Dissertation bedanken.

Mein Dank gilt ebenso Hannes Münchow und Štěpán Bahník, durch deren Einfluss ich die Wissenschaft für mich entdeckt habe. Ihre ausgewogene Mischung aus Anleitung, Unterstützung und der Freiheit, eigene Wege zu gehen, hat entscheidend dazu beigetragen, dass ich mich heute beruflich und akademisch an einem Punkt befinde, an dem ich tiefe Zufriedenheit empfinde. Mein herzlicher Dank richtet sich auch an das Team-CROCODILE und an alle meine Kolleginnen und Kollegen des Lehrstuhls für Psychologie IV. Die kompetente und engagierte Zusammenarbeit des Teams war unerlässlich für die Realisierung des Projekts. Ebenso schätze ich die Unterstützung und die positive Zeit, die ich am Lehrstuhl für Psychologie IV erleben durfte.

Abschließend gebührt ein besonderer Dank meinen Eltern, die mich mit all dem ausgestattet haben, was ich für meinen Weg benötigt habe, und deren anhaltende Unterstützung für mich unersetzlich ist. Ebenso möchte ich meiner Partnerin Charlotte tiefsten Dank aussprechen. Ihre unerschütterliche Unterstützung, besonders während der turbulenten Schlussphase dieser Arbeit, war von unschätzbarem Wert.