

KU LEUVEN

FACULTEIT PSYCHOLOGIE EN  
PEDAGOGISCHE WETENSCHAPPEN

**PULLED IN THE WRONG DIRECTION! RETHINKING  
RESPONSE CONTROL IN COMPUTATIONAL  
MODELS OF CONFLICT TASKS**

Masterproef aangeboden tot het  
verkrijgen van de graad van  
Master of Science in de  
psychologie door  
**Stien Mommen**

promotor: Dr. Luc Vermeylen  
copromotor: Prof. Kobe Desender



*Copyright informatie:*

*studiewerk van een student in het kader van de  
academische opleiding en examenbeoordeling.*

*Na deze beoordeling vond geen correctie plaats  
van het studiewerk.*



KU LEUVEN

FACULTEIT PSYCHOLOGIE EN  
PEDAGOGISCHE WETENSCHAPPEN

**PULLED IN THE WRONG DIRECTION! RETHINKING  
RESPONSE CONTROL IN COMPUTATIONAL  
MODELS OF CONFLICT TASKS**

Masterproef aangeboden tot het  
verkrijgen van de graad van  
Master of Science in de  
psychologie door  
**Stien Mommen**

promotor: Dr. Luc Vermeylen  
copromotor: Prof. Kobe Desender



## Summary

In daily life, we constantly make decisions that require us to navigate multiple sources of information. For example, when deciding whether to push or pull to open a door, the design of the door handle may trigger the automatic response tendency to pull, while a “push” sign indicates the opposite response. This mismatch illustrates how various informational cues can signal different response options that compete for activation, thereby placing demands on our ability to control responses. Such situations have sparked a large body of research to study how well individuals can flexibly control their responses during cognitive conflict.

The central problem identified and addressed in this Master Thesis is that current conceptualizations of response control lack the specificity to provide a complete picture of this control ability. Traditionally, researchers have relied on summary statistics that only provide a coarse description of response control. The move toward computational modeling, with the Diffusion Model for Conflict tasks (DMC), allows response control to be studied in terms of latent processes that give rise to behavior. Despite these advances, I argue that a critical oversight remains: current model-based measures of response control neglect the full state of the decision-maker. Specifically, distractor strength, as quantified by the model’s “peak amplitude” parameter, is used as a proxy for response control. In other words, the notion of response control is often approached in a narrow and isolated manner. To address this oversight, the present thesis aims to refine the conceptualization of response control by taking the broader decision context into account. The first part therefore investigates how peak amplitude interacts with other DMC parameters, using a simulation-based approach. Based on these findings, alternative measures of response control are developed and systematically evaluated against comparative criteria. In the final part of the present work, the most robust measure is applied on empirical data in relation to the existence of a common response control mechanism underlying performance in different conflict tasks.

Through theoretical simulations, the results show that peak amplitude should not be considered in isolation to study response control. Rather, the effect of distractor strength (peak amplitude) on behavior depends on the amount of evidence one accumulates (decision boundary) and efficiency of evidence accumulation (drift rate). This has important implications: two individuals with the same peak amplitude may nevertheless differ greatly in how cognitive conflict impacts their decision formation. Therefore, I provide a novel conceptualization of response control that accounts for the dynamic interplay between automatic and controlled activation. Specifically, Triple R emerges as the most robust model-based measure of response control among the set of candidates outlined in this thesis. Furthermore, when applied to empirical data, the significant positive correlation of Triple R between different conflict tasks provides evidence for a common response control mechanism across contexts. This finding speaks directly to the long-standing debate on the domain-generalty of response control and demonstrates the value of Triple R as a context-sensitive measure of response control.

### **Words of Gratitude**

Built on the foundations of support, this thesis has grown as much as I have. A journey of discovery, a quest of wonder and perseverance. And I am deeply grateful to have the best possible guides who helped me navigate the way. At the forefront of this journey, I would like to thank my promotor, Dr. Luc Vermeulen. Your insights served as a compass through my first steps into this fascinating world of diffusion models and response control. Thank you for creating an open and supportive atmosphere. Your enthusiasm and curiosity were truly inspiring and had a way of drawing me deeper into this project, in the best possible way.

I would also like to express my gratitude to my co-promotor, Prof. Kobe Desender. Your valuable feedback has been instrumental in refining this work. I am moreover grateful for the trust you have placed in me. Thank you for seeing my potential as a future researcher. Your support has extended far beyond the pages of this thesis, and I am deeply grateful for your help mapping out the road ahead. I cannot wait to build further on this work and continue this journey with the guidance of both of you.

Last but not least, I want to thank my mum. Even before I felt the warmth of the sun, even before I saw the light of the moon to guide me, you have always been by my side from the very start. Without your endless support, I would not be where I am today. Thank you for providing me with the safe space to develop myself and for everything you have made possible for me. I wish everyone a friend like you.

No words could truly capture what this journey has meant to me. Looking back, this thesis marks more than the end of my studies. What has started as an intriguing interest in diffusion models has grown into a concrete research direction. Bring it on, future.

### **Approach and Contribution Statement**

The origins of this thesis can be traced back to the course *Psychophysics*, taught by Prof. Kobe Desender, in which I got introduced to evidence accumulation models. In one of the journal clubs of this course, led by Dr. Luc Vermeulen, I became captivated by the way these models could be used to answer psychological questions. It was then that I knew I wanted to study decision making and diffusion models in greater depth. For this reason, I chose to submit a student-proposed thesis topic in consultation with Dr. Luc Vermeulen as promotor because of his experience in computational modeling and cognitive control.

Looking back on the journey, this thesis has been shaped by a general approach that allowed for continuous refinement. Specifically, it involved an exchange of ideas with my promotor Dr. Luc Vermeulen, which I could use as input to build on. Subsequently, I reported back on the progress I had made, for which the feedback of my promotor enabled me to refine my work again. Through these repeated stages, I could always draw on the expertise and supervision of Dr. Luc Vermeulen, while having the space to work autonomously.

In addition to this general approach, this thesis relies primarily on simulated data from the Diffusion Model for Conflict tasks. For this purpose, I wrote a custom C++ script implementing the model as specified in Ulrich et al. (2015). For the empirical application at the end of this thesis, I made use of the open-access datasets reported in Hedge et al. (2022). I performed all analyses and interpreted the results, which were discussed with Dr. Luc Vermeulen for feedback. All scripts are thus custom-made for this thesis, as are all the figures presented in this work. AI-based tools were only used in a supportive way during coding (e.g., troubleshooting). I am moreover responsible for the full writing of this thesis, refined based on the insights of my promotor Dr. Luc Vermeulen and co-promotor Prof. Kobe Desender.

## Table of Contents

|                                                                                                            |     |
|------------------------------------------------------------------------------------------------------------|-----|
| Summary .....                                                                                              | I   |
| Words of Gratitude.....                                                                                    | II  |
| Approach and Contribution Statement.....                                                                   | III |
| Table of Contents .....                                                                                    | IV  |
| List of Tables.....                                                                                        | VI  |
| List of Figures .....                                                                                      | VII |
| Pulled in the Wrong Direction! Rethinking Response Control in Computational Models of Conflict Tasks ..... | 1   |
| Interference Effects in Conflict Tasks .....                                                               | 1   |
| Diffusion Modeling of Cognitive Conflict .....                                                             | 3   |
| Unpacking the Diffusion Model for Conflict Tasks .....                                                     | 5   |
| Conceptual Problem .....                                                                                   | 8   |
| Current Goals.....                                                                                         | 11  |
| Methods.....                                                                                               | 12  |
| DMC Implementation.....                                                                                    | 12  |
| Goal 1: Moving Beyond Isolated Effects of Distractor Strength .....                                        | 14  |
| Simulation Procedure.....                                                                                  | 14  |
| Statistical Analysis.....                                                                                  | 15  |
| Goal 2: Toward a Context-Sensitive Measure of Response Control.....                                        | 16  |
| Construction of Candidate Measures .....                                                                   | 16  |
| Measure Evaluation .....                                                                                   | 17  |
| Goal 3: Revisiting Domain-Generality of Response Control.....                                              | 18  |
| Empirical Datasets .....                                                                                   | 18  |
| Statistical Analysis.....                                                                                  | 19  |
| Results .....                                                                                              | 19  |
| Goal 1: Moving Beyond Isolated Effects of Distractor Strength .....                                        | 19  |
| Amplitude-Boundary Interactions .....                                                                      | 19  |
| Amplitude-Drift Rate Interactions .....                                                                    | 22  |
| Interim Conclusion.....                                                                                    | 23  |
| Goal 2: Toward a Context-Sensitive Measure of Response Control.....                                        | 23  |
| Development of Candidate Measures .....                                                                    | 24  |

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| Peak-to-Bound Ratio (PTB-Ratio) .....                                   | 24 |
| Integrated Evidence Capture (IEC) .....                                 | 25 |
| Relative Route Regulation (Triple R) .....                              | 26 |
| Validation of Candidate Measures .....                                  | 28 |
| Inclusion of All Latent Parameters Shaping Response Control .....       | 29 |
| Meaningful Correlations with Behavioral Interference. ....              | 29 |
| Unique Mapping to Group Differences.....                                | 32 |
| Interim Conclusion.....                                                 | 36 |
| Goal 3: Revisiting Domain-Generality of Response Control .....          | 37 |
| Discussion .....                                                        | 39 |
| Distractor Strength is Only Part of the Story.....                      | 39 |
| Measuring Response Control at the Construct Level .....                 | 39 |
| Peak Amplitude as an Incomplete Measure of Response Control .....       | 40 |
| Accounting for All Parameters is Necessary but Not Sufficient .....     | 40 |
| Triple R as a Novel Context-Sensitive Measure of Response Control ..... | 41 |
| Implications for Domain-Generality .....                                | 42 |
| Limitations and Future Directions .....                                 | 43 |
| Conclusion .....                                                        | 45 |
| References .....                                                        | 46 |

### List of Tables

|                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 1: Summary of DMC Parameters .....                                                                                                                    | 14 |
| Table 2: ROC-Based Classification of High vs. Low Decision Boundary and DeLong<br>Comparisons of Response Control Measures Relative to Peak Amplitude ..... | 34 |
| Table 3: ROC-Based Classification of High vs. Low Drift Rate and DeLong Comparisons of<br>Response Control Measures Relative to Peak Amplitude.....         | 36 |
| Table 4: Overview of Response Control Measures and Evaluation Criteria .....                                                                                | 37 |

### List of Figures

|                                                                                                                                              |    |
|----------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1: Illustration of Classic Conflict Tasks .....                                                                                       | 2  |
| Figure 2: Illustration of Task-Specific Delta Slopes .....                                                                                   | 4  |
| Figure 3: Diffusion Model for Conflict Tasks (DMC).....                                                                                      | 6  |
| Figure 4: DMC Dynamics in Classic Conflict Tasks .....                                                                                       | 7  |
| Figure 5: Multi-Level Conceptual Framework .....                                                                                             | 9  |
| Figure 6: Heatmaps of Simulated Interference Effects in Accuracy and RT across<br>Combinations of Peak Amplitude and Decision Boundary ..... | 20 |
| Figure 7: Heatmaps of Simulated Interference Effects in Accuracy and RT across<br>Combinations of Peak Amplitude and Drift Rate .....        | 22 |
| Figure 8: Integrated Evidence Capture (IEC) .....                                                                                            | 26 |
| Figure 9: The Effect of Tau on Automatic Activation .....                                                                                    | 28 |
| Figure 10: Associations Between Behavioral Interference, DMC Parameters and Candidate<br>Measures.....                                       | 31 |
| Figure 11: Mapping of Response Control Measures and Error Cost for Group Differences in<br>Decision Boundary.....                            | 33 |
| Figure 12: Mapping of Response Control Measures and Error Cost for Group Differences in<br>Drift Rate.....                                   | 35 |
| Figure 13: Forest Plot of Correlations Between Task Pairs .....                                                                              | 38 |

## **Pulled in the Wrong Direction! Rethinking Response Control in Computational Models of Conflict Tasks**

In a world full of distractions – ping, a new notification on your phone – the ability to ignore irrelevant information is essential for optimal task performance (Ophir et al., 2009), educational achievements (Martin et al., 2025) and even psychological well-being (Toh & Yang, 2024). Distractions pose a big challenge to human decision making, as decisions are typically guided by multiple sources of information simultaneously. Some of these sources are directly relevant to the task at hand, while task-irrelevant sources may activate an automatic response that competes for execution. Exposure to irrelevant information may guide us away from our goal, which requires us to regain focus and navigate these distractions. For example, when traveling to the UK, visitors often tend to look left first when crossing the street. However, as cars drive on the opposite side of the road, this automatic response tendency must be suppressed to ensure safe navigation. A situation like this illustrates the need for response control, a core facet of cognitive control (Logan et al., 2014).

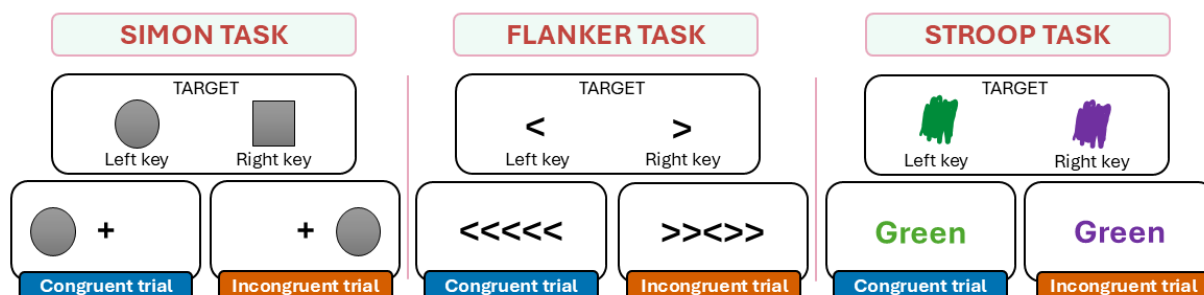
### **Interference Effects in Conflict Tasks**

Response control (or response inhibition) refers to an individual's ability to suppress an automatic, task-irrelevant response tendency and select a task-relevant response instead (Luna et al., 2010). Although this control ability enables flexible, goal-directed behavior, the suppression of task-irrelevant responses is not flawless. In fact, performance costs are typically observed in conflict tasks, where the stimulus dimension relevant to the task at hand must be filtered out from task-irrelevant features. For example, in a visual variant of the Simon task (Simon, 1969), responses are typically slower and more error-prone when spatial location conflicts with the correct response based on stimulus shape (see Figure 1). Another classic conflict task in which performance can be influenced by task-irrelevant information is the Flanker task (Eriksen & Eriksen, 1974). In this task, individuals are instructed to respond to a central target while ignoring the surrounding flankers. As illustrated in Figure 1, the target can face left or right, with responses assigned to the left and right response key respectively. When the flankers match the target, both sources of information activate the same response, facilitating performance. However, when the central target is surrounded by flankers pointing in the opposite direction, different responses compete for activation, resulting in behavioral interference. A similar influence of distracting information on task performance is observed in the Stroop task (Stroop, 1935). This classic conflict task requires naming the ink color of a color word, that either matches (congruent) or conflicts (incongruent) with the word's semantic

meaning. As shown in Figure 1, when the word “green” is presented in purple ink, the irrelevant semantic word dimension activates an automatic response that competes with the correct response based on the purple ink color.

**Figure 1**

*Illustration of Classic Conflict Tasks*



*Note.* Exemplary classic conflict paradigms in which participants are instructed to respond to a designated target feature using the corresponding response key while ignoring task-irrelevant information.

Performance in conflict tasks is commonly evaluated by comparing mean accuracy and reaction time (RT) between congruent and incongruent trials, the so-called interference effects. On congruent trials, task-relevant and task-irrelevant information signal the same response, whereas on incongruent trials, the two sources of information indicate different responses. As a result, incongruent trials are characterized by lower average accuracy and longer mean RTs compared to congruent trials. These interference effects have traditionally been used as an index of how well individuals can control their responses during cognitive conflict. Stronger control over the erroneous response tendency is assumed to reduce the performance cost associated with incongruent trials, resulting in smaller interference effects. For instance, a smaller interference effect for bilinguals compared to monolinguals in a Simon task has been interpreted as more inhibitory efficiency in bilingual participants (Bialystok et al., 2004). The finding that patients diagnosed with mild cognitive impairment showed increased interference effects in a Flanker task has been taken as evidence for less efficient response inhibition compared to a group of healthy controls (Wylie et al., 2007). Similarly, larger Stroop effects observed in alcohol-dependent individuals have been attributed to a deficit in inhibiting prepotent responses (Noël et al., 2013). However, such summary statistics of behavior only provide a coarse and impure description of response control. Specifically, since interference effects do not offer insights into the processes that give rise to observed performance (Verbruggen et al., 2014),

they cannot explain why differences in response control are observed. Moreover, interference effects do not necessarily indicate differences in true response control ability, as they reflect an unknown mixture of various underlying processes (White et al., 2018). Luckily, advances in computational modeling have provided us with tools that help to examine the underlying processes giving rise to behavioral performance. These models can furthermore quantify latent constructs and relate them to individual differences that may otherwise be looked over at by the conventional approach of interpreting interference effects in isolation (Farell & Lewandowsky, 2015).

### **Diffusion Modeling of Cognitive Conflict**

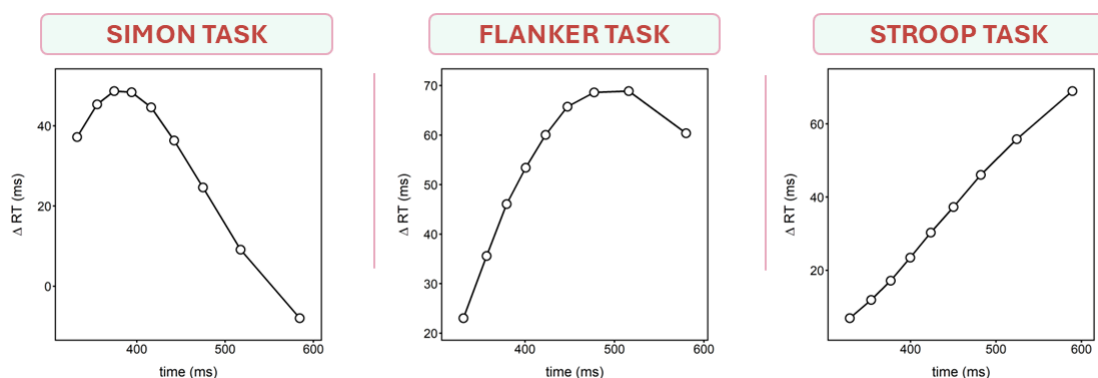
A prominent class of decision-making models conceptualizes the decision process as the continuous accumulation of noisy evidence over time, until a decision boundary is reached. This class of computational models is popularized by the Drift Diffusion Model (DDM; Ratcliff, 1978). In this framework, the decision process starts at an intermediate point between two decision boundaries. In two-choice decisions, such as in conflict tasks, the upper boundary typically corresponds to the correct response option, whereas the lower boundary indicates the incorrect response. The decision process ends when the evidence accumulation trajectory reaches one of those boundaries, at which the decision-maker commits to a response option. As a core feature of diffusion models, task performance can be explained in terms of a set of parameters corresponding to the latent processes giving rise to observed behavior. Specifically, the setting of decision boundary indicates the amount of evidence that is accumulated before committing to a response option, while the drift rate parameter captures the efficiency of evidence accumulation. For example, by applying the diffusion framework, researchers were able to identify a decrease in drift rate as one of the underlying processes of the decrement effects of alcohol on perceptual decision making (van Ravenzwaaij et al., 2012). Furthermore, processes occurring before and after the actual decision process, such as time needed to encode the stimulus, are accounted for by non-decision time. The importance of decomposing RT in non-decision time and the time needed to make the decision itself is illustrated by studies on aging. For example, a higher decision boundary setting, reflecting more response caution, is one of the underlying factors related to longer RTs in older compared to younger adults, next to slowing down in non-decision processes (e.g., Kelber et al., 2025; Servant & Evans, 2020).

The widespread use of the diffusion framework is reflected in the large number of studies analyzing conflict data with the “vanilla” DDM (e.g., Fröber & Lerche, 2023; Haller et al., 2021; Westfall et al., 2018). Although the vanilla DDM has been successfully applied across

a broad range of research domains (for a review, see Ratcliff & McKoon, 2008), its application to decision making during cognitive conflict is more complex. The vanilla DDM assumes a constant stream of evidence feeding into the decision process, an assumption that is clearly violated in conflict tasks where distracting information (e.g., flankers) exerts a devastating effect relatively early on but attenuates later during a trial. As a result, the vanilla DDM describes empirical data from conflict tasks poorly (Servant et al., 2014). The extent to which (diffusion) models can account for empirical data is commonly evaluated using full RT distributions. These distributions reveal distinct patterns when examining fast vs. slow decisions in different contexts. For example, examination of interference effects across the entire RT distribution in a Flanker task revealed that patients with mild cognitive impairment experienced more interference for slower decisions (Wylie et al., 2007). More generally, when cognitive conflict emerges relatively late during the decision process, behavioral interference is typically larger for slower compared to fast decisions. In contrast, when conflict is strongest early during decision formation, interference tends to be larger for fast responses. These task-specific patterns are captured by delta plots (De Jong et al., 1994), which represent quantile differences between congruent and incongruent RT distributions of correct responses. As illustrated in Figure 2, the classic conflict tasks show characteristic delta slope patterns. Positive-going delta slopes, indicating larger interference effects for longer responses, are often found in Stroop (Pratte et al., 2010) and Flanker tasks (Van Den Wildenberg et al., 2010). In contrast, the analysis of Simon task distributions typically reveals decreasing interference effects for slower responses, as reflected in negative-going delta slopes (Pratte et al., 2010).

**Figure 2**

*Illustration of Task-Specific Delta Slopes*

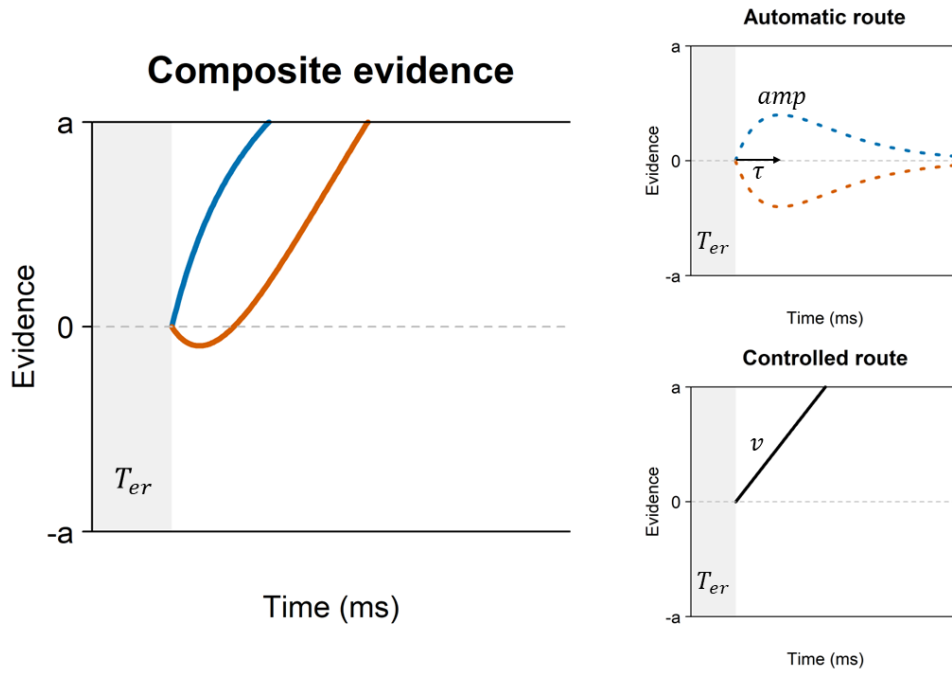


*Note.* Exemplary delta plots obtained from simulations with the DMC, illustrating the difference in mean RT between incongruent and congruent trials (y-axis) as a function of mean RT for each bin (percentiles 0.1-0.9) averaged across congruency conditions (x-axis).

Negative-going delta slopes have posed a challenge for many computational models. In particular, neither the vanilla DDM nor more recent diffusion models, such as the Dual-Stage Two-Phase model (Hübner et al., 2010) and the Shrinking Spotlight model (White et al., 2011), are able to explain negative-going delta slopes (Ulrich et al., 2015). To address this shortcoming, the Diffusion Model for Conflict tasks (DMC; Ulrich et al., 2015) was specifically developed with the aim to explain the distinct distributional patterns in different conflict tasks. Crucially, as shown in Figure 2, the DMC can account for both positive and negative-going delta slopes. As a result, the DMC overcomes a key limitation of earlier models and offers a wide lens to study how individuals control their responses in different conflict tasks.

### Unpacking the Diffusion Model for Conflict Tasks

The Diffusion Model for Conflict tasks (DMC) has been introduced as an extension of the vanilla DDM by formally incorporating the conceptual distinction of a controlled and automatic route to model conflict data (Ulrich et al., 2015). The controlled route processes task-relevant information, while irrelevant information, such as the distractors in a Flanker task, is processed by the automatic route. Together, the input of both routes is combined into a single evidence accumulation process until a decision boundary ( $\pm a$ ) is reached (see Figure 3). The controlled route in the DMC can be considered a reflection of the evidence accumulation process as modeled in the vanilla DDM. Specifically, activation of the controlled route, driven by task-relevant information, is assumed to increase linearly. The slope of this linear function is quantified by the time-independent drift rate parameter ( $v$ ), indicating how efficiently evidence accumulation goes. Within the framework of the DMC, the drift rate of the controlled route is assumed to be larger than zero, such that accumulated information relevant to the task at hand provides evidence for the correct response option. As illustrated in Figure 3, the controlled route shows a gradual increase over time, while the influence of automatic activation is assumed to be rather short-living. The automatic route, driven by task-irrelevant information, follows a pulse-like form that initially increases until it reaches its maximum, for which it decays back to zero again. The maximum automatic activation is quantified by the amplitude parameter ( $amp$ ), such that higher values correspond to increased distractor strength. The distinction between the controlled and automatic route helps to explain why performance is typically impaired on incongruent trials. As shown in Figure 3, both routes converge on the same response on congruent trials, whereas competing responses are activated on incongruent trials, resulting in behavioral interference.

**Figure 3***Diffusion Model for Conflict Tasks (DMC)*

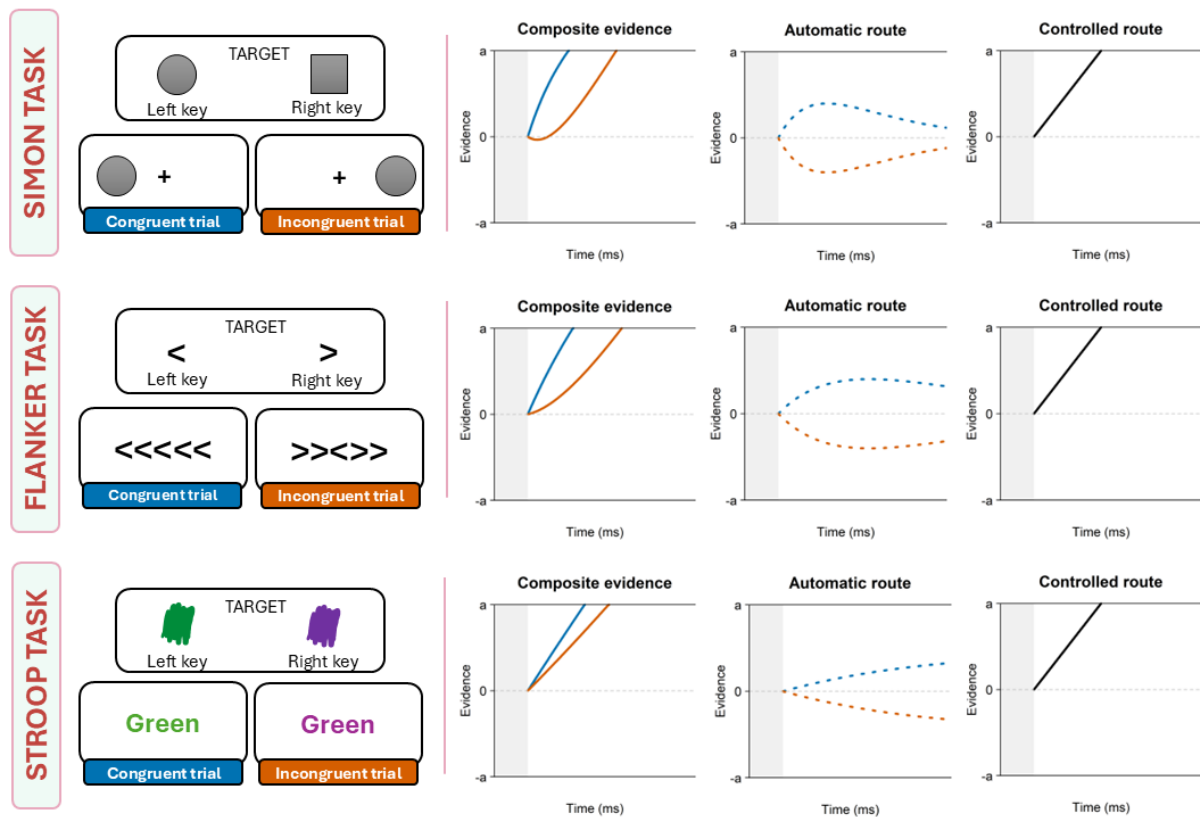
*Note.* Evidence accumulates toward one of the two decision boundaries ( $\pm a$ ), with the grey region indicating non-decision time ( $T_{er}$ ). Processing of task-relevant information along the controlled route is modeled as a linear function, for which drift rate ( $v$ ) quantifies the efficiency of evidence accumulation. In parallel, maximal automatic activation, driven by task-irrelevant information, is captured by peak amplitude ( $amp$ ) that reflects distractor strength. The time course of automatic activation is in turn indicated by the time-to-peak parameter ( $\tau$ ). Independent of the condition (blue = congruent; orange = incongruent), the controlled route guides toward the correct response boundary ( $a$ ). Similarly, on congruent trials ( $amp > 0$ ), the automatic route points to the same (correct) response boundary ( $a$ ). However, on incongruent trials ( $amp < 0$ ), the automatic route pulls to the opposite response boundary ( $-a$ ), resulting in cognitive conflict. The final decision, and thus response activation, is based on the superimposed activations of both the controlled and automatic route as reflected in the composite evidence traces.

One of the key features of the DMC's mathematical architecture is the time-to-peak parameter ( $\tau$ ) of the automatic route, which specifies when automatic activation peaks. This parameter allows the model to account for conflict data of various conflict tasks that differ in time course of automatic activation. As can be seen in Figure 4, the maximum peak of automatic activation is typically reached earlier in a Simon task than in a Flanker task (Mittelstädt et al.,

2022). In turn, estimates of the time-to-peak parameter are generally larger in a Stroop compared to a Simon and Flanker task (Ellinghaus et al., 2024). These differences reflect temporal variation in the source of cognitive conflict (Janczyk et al., 2025). For example, the Stroop task involves a semantic component, as the meaning of a color word automatically activates a competing response that interferes with ink color. The activation of semantic representations typically takes more time than the activation of automatic action representations (Liotti et al., 2000; Valle-Inclán, 1996). Importantly, it is this flexibility in the time course of automatic activation, enabled by the time-to-peak parameter ( $\tau$ ), that underlies the ability of the DMC to account for both positive and negative delta-slopes (cf. Figure 2). Furthermore, the architecture of the DMC describes a gradual decline of distractor-driven automatic activation over time. Unlike models that assume a form of all-or-nothing processing of distractors, such as the Dual-Stage Two-Phase model (Hübner et al., 2010), the DMC implements decline of automatic activation within a continuous evidence accumulation framework. In sum, the DMC has important features to advance our understanding of response control.

**Figure 4**

*DMC Dynamics in Classic Conflict Tasks*



*Note.* Exemplary DMC dynamics in Simon, Flanker and Stroop tasks are illustrated using simulated model trajectories, for which only the time-to-peak parameter differs (Simon  $\tau = 72$ ;

Flanker  $\tau = 135$ ; Stroop  $\tau = 505$ ). In other words, the time at which the automatic activation peaks (i.e., when distractor influence is most pronounced) varies across conflict tasks. As a result, differences in time course are also apparent in composite evidence trajectories. Blue traces represent congruent trials; orange lines illustrate incongruent trials.

### Conceptual Problem

Having outlined how the DMC's architecture allows to model decision making in situations of cognitive conflict, an important question remains: how can one's true response control ability be measured within this framework? As illustrated in Figure 5, response control can be considered a latent construct that manifests itself in behavior. In this regard, interference effects have traditionally been used as an index of response control. However, as previously noted, observed mean RT and accuracy differences between (groups of) individuals capture variance in behavior that is not necessarily related to differences in true response control ability. Rather than inferring response control directly from interference effects, advances in computational modeling have made it possible to move to the process level. Computational models, such as the DMC, serve as an essential tool, as they allow to make inferences about response control based on the parameters that capture the underlying processes giving rise to observed behavior. Importantly, this shift to the process level has gone in hand with a tendency to approach response control from an isolated perspective. In particular, in current literature, response control is often approached through indices of automatic activation elicited by task-irrelevant information, with a predominant focus on peak amplitude that captures distractor strength. For example, Hedge et al. (2022), who studied the domain-general nature of response control across different types of conflict tasks, have the following view on the relation of DMC parameters and response control:

The ability to overcome conflict is implicit in the degree of susceptibility to prepotent response activation (the amplitude it reaches), and the speed at which automatic activation peaks and is removed/decays. (p.1450)

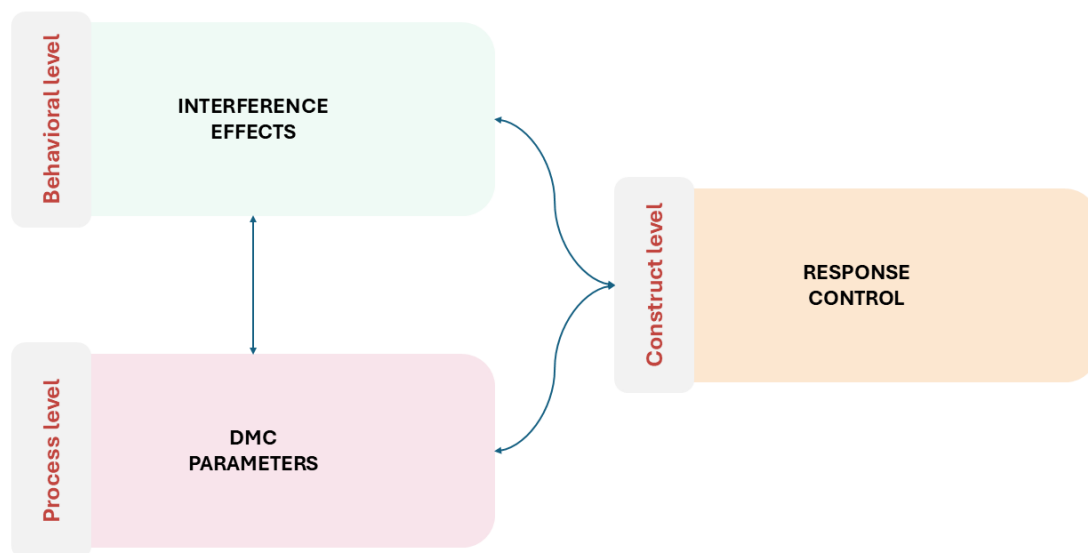
Individuals with more efficient inhibition would be expected to have either a lower amplitude and/or shorter time to peak as these are the parameters that should capture individual differences in conflict processing. (p.1450)

The notion that response control can be directly quantified by simply looking at the strength of (peak) automatic activation, is seen across a broader body of literature. For example,

studies on age-related differences in response control within the DMC framework generally assume that specific age-related inhibition deficits are reflected in larger peak amplitude and/or longer peak latencies of automatic activation (Kelber et al., 2025; Servant & Evans, 2020). Similarly, Luo et al. (2024) interpret prior findings of increased peak amplitude and time-to-peak of automatic activation following congruent trials (Koob et al., 2023; C. Luo & Proctor, 2022) as a reflection of a reduction in control. Intuitively, these (and other) researchers thus simply assume that weaker distractor strength (i.e., smaller peak amplitude of automatic activation driven by distractors) implies greater response control. However, in the current thesis I will argue and demonstrate that this assumption is too simplistic, as it overlooks complex interactions between parameter settings.

**Figure 5**

*Multi-Level Conceptual Framework*



*Note.* The conceptual problem identified and addressed in this thesis is situated at the junction of the process level and construct level. The construct of response control is often approached from an isolated process perspective, with a predominant focus on the strength of automatic activation as quantified by the DMC “peak amplitude” parameter. However, since this narrow view does not consider the full state of the decision-maker, conceptual refinement is needed to clarify how DMC parameters, and thus the underlying processes of behavior, relate to the latent construct of response control.

In previous work, Vermeylen et al. (2025) point to possibly misleading results concerning the measurement of conflict when operationalized as peak amplitude while disregarding the height of the decision boundary. For example, imagine two individuals: person

X has a peak amplitude of 20, while person Y has an amplitude of 30. Without any additional information, one might conclude that person X exhibits greater response control, as they experience less automatic activation from task-irrelevant information. However, this conclusion is incomplete without considering the setting of the decision boundary. If person X sets a low boundary of 30, their estimated peak amplitude of 20 is proportionally very high, bringing them quite close to an erroneous response. In contrast, if person Y is a more cautious decision-maker and accumulates more evidence before deciding, their larger peak amplitude is proportionally much smaller relative to a boundary setting of 60. As a result, person Y is actually less likely to make an error, despite a larger absolute automatic activation. This implies that the strength of automatic activation, which is driven by task-irrelevant information, may have a smaller impact on behavior relative to a high decision boundary. In other words, when boundary setting is high, more evidence is accumulated before a decision is made, leaving more opportunity for additional evidence to possibly counteract the effect of distractors. This suggests that someone who accumulates more evidence upon deciding (i.e., higher decision boundary) can thereby reduce the influence of distractors and thus exhibits more effective response control.

A similar way of reasoning applies when considering the drift rate. Specifically, the drift rate of the controlled route reflects the degree of processing efficiency (Weigard et al., 2025). In other words, this parameter can be considered an index of how efficiently individuals can extract task-relevant information while ignoring irrelevant features. This suggests a modulating influence of drift rate on the impact of distractor strength (i.e., peak amplitude) during decision formation, and thus a way to control conflicting responses. Even a relatively low peak amplitude could be disruptive depending on how efficiently task-relevant information is extracted (i.e., drift rate). That is, when evidence accumulation along the controlled route goes relatively slow, even minor levels of peak amplitude can interfere substantially with behavior. Conversely, when drift rate is very high, evidence accumulation to the correct response option goes efficiently somewhat regardless of the value of peak amplitude. As a result, it is more likely that the pull from distracting information toward the incorrect response is counteracted by strong controlled activation. Moreover, since a higher decision boundary is associated with slower responses (Ratcliff et al., 2015), this implies that the setting of decision boundary can modulate the time for controlled processes to act on task-relevant information. Altogether, instead of interpreting a smaller peak amplitude in isolation as an index of more effective response control (Schaefer et al., 2026), these considerations suggest that the degree of control may still differ depending on the setting of the decision boundary and the magnitude of drift rate.

The preceding argument of this thesis suggests that response control as a construct cannot be directly inferred from isolated parameters. In light of these considerations, a direct mapping between isolated parameters measured at the process level and response control at the construct level may provide an incomplete picture of response control. This has profound implications for the field as it bears directly on the inferences that can be made about response control. Most prominently, this has important implications for the long-standing debate regarding the domain-general nature of response control. Domain-general nature of response control implies that individuals who display greater response control in one task are also expected to show higher levels of response control in another context of cognitive conflict. A substantial body of research has therefore studied whether a common response control trait underlies observed behavior in different conflict tasks. In line with the traditional approach of using interference effects as a direct index of response control, the question of domain-general nature of response control has been typically investigated using correlations in interference effects between conflict tasks (e.g., Gerlach & Poirel, 2026; Khng & Lee, 2014; Rey-Mermet et al., 2018). However, when response control is measured as interference effects, the correlations between interference effects may be much smaller or larger than the correlations between the latent construct they are assumed to measure (Miller & Ulrich, 2013). Moreover, as discussed earlier, interference effects are contaminated by various factors that are not necessarily specific to the construct of interest. This highlights the need for a cognitive modelling approach to more accurately capture true response control ability. Therefore, Hedge et al. (2022) applied the DMC to examine the correlations in model parameters between datasets of Simon, Flanker and Stroop tasks. Based on the absence of between-task correlations of peak amplitude in isolation, this influential study questioned the existence of a common response control system underlying performance across different conflict tasks. However, evaluating parameter correlations in isolation does not take into account the full state of the decision-maker and thus neglects the context in which a given peak amplitude occurs. This is especially important when comparing response control within or across different conflict tasks. In sum, this raises the question of how meaningful it is to study the domain-general nature of response control, and the construct in itself when it is operationalized primarily in terms of peak amplitude parameter.

## **Current Goals**

Based on the previous discussion, it is clear that the field currently lacks a nuanced approach to quantify response control that takes into account the full state of the decision-maker. Put differently, a novel approach is needed that moves beyond an isolated process

perspective. In order to work toward a novel quantification of response control at the construct level, it is necessary to investigate the effects of peak automatic activation in relation to both the setting of the decision boundary and the efficiency at which task-relevant information is processed along the controlled route. Therefore, the first part of this thesis will study (1) amplitude-boundary interactions and (2) amplitude-drift rate interactions. The key aim of examining these in tandem is to determine to what extent the behavioral expression of cognitive conflict depends on the joint configuration of these DMC parameters, rather than on peak amplitude in isolation. Investigating such interactions is necessary to evaluate the predominant isolated approach and to get nuanced insights into how the latent processes jointly shape performance during cognitive conflict. By nuancing how the interplay of these parameters shapes the extent to which one can manage cognitive conflict, this step will provide deeper insights to work toward alternative measures that capture response control in all its facets. Accordingly, the second aim of the current work is to develop and systematically evaluate candidate model-based measures against comparative criteria to move toward a context-sensitive conceptualization of response control. Specifically, by leveraging the interactions between DMC parameters at the process level, the second section formulates novel measures to capture response control at the construct level. This step will provide a crucial shift to approach the notion of response control from isolated model parameters to an integrated quantification of the psychological construct of response control. Finally, the third aim of this thesis is to shed new light on the on-going debate regarding the domain-general nature of response control. As discussed above, while previous research has suggested that response control is uncorrelated across different conflict tasks, this research (Hedge et al., 2022) assumed a simplistic mapping between isolated parameters at the process level and response control. Therefore, using my novel quantification of response control at the construct level, I will re-analyze the data from Hedge and colleagues to approach the question of domain-generalizability through a context-sensitive framework that captures response control in all its facets.

## Methods

### DMC Implementation

In the current thesis, the Diffusion Model for Conflict tasks (DMC) was programmed in a custom script, written in C++ to enable efficient large-scale data simulations. The implementation followed the original formulation of the model by Ulrich et al. (2015). As in the general diffusion framework, RT on a given trial can be broken down into the time needed to make the decision itself and non-decision time. Non-decision time ( $T_{er}$ ) accounts for the

processes that occur before and after the decision process itself, such as the time needed to move your finger and execute a response. In the implementation of the DMC in the present work, non-decision time was sampled from a normal distribution:

$$T_{er} \sim \mathcal{N}(\mu_{T_{er}}, \sigma_{T_{er}})$$

with mean  $\mu_{T_{er}} = 300$  (ms) and  $\sigma_{T_{er}} = 20$  (ms), in line with realistic ranges reported in literature (Matzke & Wagenmakers, 2009; Hedge et al., 2022). Note that the predictions of the DMC, such as the patterns of delta plots, do not vary with properties of non-decision time (Ulrich et al., 2015). As a result, the specific choices regarding non-decision time do not affect the results presented in this thesis. This follows from the assumption within the diffusion framework that non-decision time is an additive component to the actual decision time.

Within the diffusion framework, the decision process itself is modeled as the accumulation of noisy evidence over time, in discrete steps of  $\Delta t = 1$  (ms) until a decision boundary is reached. On each trial, the decision variable starts in-between the upper decision boundary  $a$  and lower decision boundary  $-a$ . In the present implementation, the possibility of starting point variability across trials was implemented using a symmetrical beta distribution with shape parameter  $\beta = 2$ , defined on the upper and lower decision boundaries as limits and centered at midpoint zero. This provides flexibility to account for noise in the starting point of the evidence accumulation process, where in some trials, starting point is shifted closer to one decision boundary relative to the other (Ratcliff, 2013). The exact updating of the decision variable  $X$  at time  $t$  is given by:

$$X(t + \Delta t) = X(t) + \mu(t)\Delta t + s\sqrt{\Delta t}\varepsilon$$

where  $s = 4$  indicates the diffusion constant (noise scaling factor) and  $\varepsilon \sim \mathcal{N}(0, 1)$  represents noise within the evidence accumulation process, sampled from a standard normal distribution. Updating of the decision variable  $X$  is driven by  $\mu(t)$ , which reflects the superimposed activation of the automatic and controlled route. Composite evidence accumulated at time  $t$  is the combined input of the automatic route and the controlled route with drift rate  $v$ :

$$\mu(t) = amp \cdot e^{-t/\tau} \left[ \frac{t \cdot e}{(\alpha - 1)\tau} \right]^{\alpha-1} \cdot \left( \frac{\alpha - 1}{t} - \frac{1}{\tau} \right) + v$$

where  $amp$  represents the amplitude of automatic activation, which is positive for congruent trials and negative for incongruent trials. The shape parameter  $\alpha$  of the automatic route was

fixed to  $\alpha = 2$  in the present work, aligning with the original formation of the DMC. Under this parameter specification, the latency parameter  $\tau$  can be interpreted directly as the time at which the automatic activation reaches its peak. A smaller  $\tau$  implies that the automatic influence of task-irrelevant information peaks early in a trial, whereas larger values indicate a later time to peak of automatic activation. Moreover, a larger  $\tau$  broadens the course of automatic activation over a longer time window, thereby modulating cognitive conflict. By varying  $\tau$ , the model can simulate distinct temporal dynamics of automatic activation across different conflict tasks.

**Table 1**

*Summary of DMC Parameters*

| Parameter      | Description                                    |
|----------------|------------------------------------------------|
| $\mu_{Ter}$    | Mean non-decision time                         |
| $\sigma_{Ter}$ | Standard deviation of non-decision time        |
| $\beta$        | Shape parameter of starting point distribution |
| $a$            | Upper decision boundary ( $-a$ lower boundary) |
| $v$            | Drift rate of controlled activation            |
| $amp$          | Amplitude of automatic activation              |
| $\alpha$       | Shape parameter alpha of automatic activation  |
| $\tau$         | Time-to-peak of automatic activation           |

*Note.* Parameter symbols may vary across studies. Most symbols used in the present work match those in the original publication of the DMC (Ulrich et al., 2015), while others were chosen to be more meaningful within this thesis. Specific parameter values used for the analyses in the current work are reported separately for each goal in the corresponding sections below.

## **Goal 1: Moving Beyond Isolated Effects of Distractor Strength**

### ***Simulation Procedure***

In order to investigate the effects of distracting information in relation to (1) the setting of decision boundary and (2) the efficiency of task-relevant evidence accumulation, a simulation-based approach was applied. In the first analysis, peak amplitude and decision boundary were systematically varied across predefined ranges to study how changes in each parameter affect behavioral interference. Boundary values ranged from 10 to 110 in increments of 2, while amplitude values ranged from 0 to 100 in increments of 2. For each combination within this two-dimensional parameter space, 5000 trials per parameter combination were

simulated using the DMC as described above, with a default drift rate of 0.4 and a time-to-peak parameter of  $\tau = 50$ , approximating the time course of automatic activation in a Simon task. Behavioral outcome indices related to RT and accuracy were then computed for each combination of peak amplitude and decision boundary. Because response control is a latent construct that shapes observed behavior (cf. Figure 5), behavioral indices were mapped as a third variable onto parameter space. Resulting patterns were visualized as heatmaps, with color gradients reflecting the magnitude of each behavioral index across parameter space. The same simulation-based approach was applied to study the effects of changes in peak amplitude alongside changes in drift rate. This allows to identify how different combinations of amplitude and drift rate shape the degree of interference one experiences. In this second analysis, drift rate values ranged from 0.1 to 1.2 in increments of 0.1 while peak amplitude was systematically varied from 0 to 100 in steps of 2, as in the previous analysis. Given that interference effects stay relatively stable for varying drift rates when time-to-peak of automatic activation occurs early (Janczyk et al., 2025), data was simulated for  $\tau = 135$ , which mimics the time course of a Flanker task. The decision boundary was fixed at  $a = 60$ , representing a realistic default setting in a Flanker task (Mittelstädt et al., 2022).

### *Statistical Analysis*

Next to visual inspection of the generated heatmaps, I evaluated the patterns in the resulting heatmaps more formally by regression analyses in R. Because these analyses are based on simulated data with a large number of trials per parameter combination, it is important to reflect on the influence of sample size on p-values. Although it is often assumed that sufficiently large sample sizes will inevitably yield statistically significant results, this comes with an important caveat. Only when a true effect exists, p-values become increasingly smaller with larger sample size. However, when  $\mathcal{H}_0 = \text{true}$  (no effect), increasing sample size does not affect p-values which are theoretically still expected to be uniformly distributed (Hofmann, 2016). Put differently, statistical testing remains appropriate with a large number of simulated data, especially since the regression analyses in the present work are used to formally support the visual analysis of the patterns in the heatmaps. Therefore, separate multiple regression analyses were conducted for the simulated interference effects in both accuracy and RT. Specifically, first investigating the interplay of peak amplitude and decision boundary, interference effects were modeled using peak amplitude, decision boundary and their interaction as predictors. To capture potential non-linear dynamics in parameter space, this model was compared against a quadratic model that included second-order orthogonal polynomial terms for both predictors

alongside their full interactions. Model fit was evaluated using the Bayesian Information Criterion (BIC). This criterion considers goodness-of-fit, while penalizing for model complexity. If a lower value of BIC is observed for the quadratic model relative to the linear model, this indicates that the quadratic model provides a better fit of the data. For the selected model, the significance of the predictors was assessed using omnibus F-tests derived from the `anova()` function in R, testing the combined contribution of linear and quadratic components for each term. The same regression analysis was applied to investigate the interplay of peak amplitude and drift rate on behavioral interference, in which decision boundary was replaced by drift rate as a predictor. This procedure allowed to consider the effects of distractor strength (i.e., peak amplitude) within the full state of the decision-maker (i.e., the interplay of parameters) more formally.

## **Goal 2: Toward a Context-Sensitive Measure of Response Control**

### ***Construction of Candidate Measures***

To work toward a measure of response control that takes the broader decision-context into account, I developed and systematically evaluated three candidate measures designed to capture response control at the construct level. Because the development of these candidate measures was based on the results of the first section and constituted one of the main aims of the current work, their formal construction and definitions are presented in the corresponding results section of Goal 2. The present section therefore provides only a brief rationale of their development. Specifically, each measure was constructed to shift the focus away from a direct mapping between isolated parameters and response control, toward an integrated quantification of the psychological construct. Although all candidate measures were developed with the same goal in mind, I approached each of them from a slightly different angle. The first candidate was originally introduced by Vermeulen et al. (2025) as a measure of sensitivity to task-irrelevant information. The current work re-approached this measure from a broader perspective to examine how well it translates into a measure of response control at the construct level itself. As a second candidate measure, I developed a novel measure of response control based on the composite evidence accumulation trajectories. This way, the dynamic properties of the unfolding decision process are taken into account. Finally, I constructed a third candidate measure based on the joint contribution of parameters along the automatic and controlled route. The findings of the first part of the current study formed the basis of this third alternative quantification, accounting for the directional influence in which these parameters shape how well someone can control their responses during cognitive conflict.

### ***Measure Evaluation***

To enable a comparative evaluation of peak amplitude as a stand-alone measure and the alternative candidate measures presented in the current work, these measures of response control were evaluated against three criteria. The first criterion comprised the extent to which each measure accounts for the broader decision-context. Evaluation of this criterion is based on a conceptual inspection of whether the measures capture the contribution of each parameter that jointly shape the degree of response control.

Complementing this conceptual evaluation, the second criterion assessed whether the measures are meaningfully related to behavior. While a measure of a latent construct may be theoretically sound, it remains important to evaluate how it relates to task performance as an observable, but impure, reflection of the latent construct. Therefore, I simulated data with variability in the latent parameters that give rise to variability in task performance. Simulated datasets were generated for 1000 individuals, with DMC parameters sampled from normal distributions for each individual. Specifically, peak amplitude ( $\mu = 27.5, \sigma = 4.17$ ), decision boundary ( $\mu = 62.5, \sigma = 5.83$ ), drift rate ( $\mu = 0.5, \sigma = 0.1$ ), starting point variability ( $\mu = 2.5, \sigma = 0.167$ ) and non-decision time ( $\mu = 335, \sigma = 21.67$ ). The selected parameter ranges align with those reported in Hedge et al. (2022) to mimic individual differences in model parameters governing the decision process. Datasets were generated for the three classic conflict tasks that differ in temporal dynamics of automatic activation, with 10 000 trials per task. Specifically, the time-to-peak parameter ( $\tau$ ) was sampled from task-specific normal distributions with  $\mu = 72$  (Simon),  $\mu = 135$  (Flanker) and  $\mu = 505$  (Stroop), each with  $\sigma = 16.67$ . For each participant, their degree of response control was quantified using the candidate measures, alongside computation of behavioral interference effects for both RT and accuracy as a reflection of task performance. The extent to which each measure captures variability in behavioral interference is quantified using Pearson correlations.

Next to being meaningfully tied to behavior, a valid measure of response control should appropriately map differences in task performance onto differences in the latent construct. For example, individuals who show better task performance due to a higher decision boundary setting should be identified by the measure as individuals showing greater control of responses than those who perform worse because of lower boundaries. If the measure assigns similar values to the construct despite differences in performance related to decision boundary, this indicates that the measure cannot explain performance variation related to the specific component (i.e., decision boundary) that shapes response control. Therefore, I simulated data with the DMC for two groups, each consisting of 250 individuals, with low boundary (i.e.,

impulsive decision-makers) vs. high decision boundary setting (i.e., cautious decision-makers). For each individual, decision boundary settings were sampled from normal distributions with group-specific means: low boundary ( $\mu = 60, \sigma = 5$ ) vs. high boundary ( $\mu = 120, \sigma = 5$ ). Peak amplitude ( $\mu = 40, \sigma = 5$ ) and drift rate ( $\mu = 0.5, \sigma = 0.01$ ) were sampled from the same normal distributions in both groups. As in the data simulation procedure for criterion 2, data were generated for the same three conflict tasks (10 000 trials per task). Specifically, the time-to-peak parameter ( $\tau$ ) was sampled from normal distributions (Simon:  $\mu = 72$ , Flanker  $\mu = 135$ , Stroop:  $\mu = 505$ , all  $\sigma = 16.67$ ), thereby preserving the distinct temporal dynamics of automatic activation across tasks as in criterion 2. Beyond differences in decision boundary, studies have also reported differences between groups in processing of task-relevant information. Therefore, I applied the same simulation approach to evaluate how the different measures of response control map onto group differences in low vs. high drift rate. Two groups of 250 individuals were simulated, with drift rate values sampled from normal distributions with group specific means: low drift rate ( $\mu = 0.2, \sigma = 0.01$ ) vs. high drift rate ( $\mu = 0.8, \sigma = 0.01$ ). All simulation settings and other parameters, except for boundary ( $\mu = 90, \sigma = 5$ ), were identical to those described above. To support visual analysis and to quantify the ability of each measure to discriminate between groups, Receiver Operating Characteristic (ROC) analyses were conducted. For each conflict task, ROC curves were computed using group membership (i.e., low vs. high decision boundary/drift rate) as the binary outcome and each measure of response control as the predictor. Classification performance of each measure was quantified using the area under the ROC curve (AUC), with values of 0.50 indicating chance-level classification while values approaching 1 indicating near-perfect discrimination between the groups. Each candidate measure was compared against peak amplitude within each task using DeLong tests. Positive z-values would thus indicate better classification performance of the alternative measure. Altogether, this procedure enables to evaluate whether a measure intended to capture response control is sensitive to group differences related to the component(s) that shape the psychological construct.

### **Goal 3: Revisiting Domain-Generality of Response Control**

#### ***Empirical Datasets***

To shed new light on the debate regarding the domain-generality of response control, I revisited previous work by Hedge et al. (2022) questioning the existence of a common response control mechanism underlying performance in different conflict tasks. Since this conclusion was based on (non-significant) single parameter correlations between tasks, I reconsidered this

question with my novel, context-sensitive measure of response control. Specifically, I reanalyzed seven empirical datasets comprising different variants of Simon, Flanker and Stroop tasks by applying a multilevel meta-analytic approach. The fitted DMC parameters reported by Hedge et al. (2022) formed the primary input for the current analysis.

### ***Statistical Analysis***

Since the original analysis script by Hedge and colleagues was not publicly available, I reconstructed the procedure based on the methodological description in their publication, combined with additional information listed in the manual to conduct a meta-analysis in R with the metafor package (Viechtbauer, 2010). I computed Spearman's rho correlations for each individual parameter between task pairs, within each dataset. This resulted in a set of correlation values, where parameters from one task were correlated with their counterparts in the paired task (e.g., Flanker boundary correlated with Simon boundary). This procedure was applied systematically to all DMC parameters as well as to my novel quantification of response control. Given that random-effects meta-analyses of raw correlations can lead to biased meta-analytic estimates (Stanley et al., 2024), I applied Fisher's z-transformation to circumvent that standard errors of correlations coefficients depend on the size of the coefficient. These transformed values formed the input of a multilevel random effects model. To account for dependencies in the data, the model included a random effects structure with two levels of dataset and task pair. This accounts for the possibility that correlations obtained within the same dataset may be more similar to each other than those coming from different datasets, as well as correlation values between specific task pairs may be higher between some than others because of more similar task structures. Finally, the meta-analytic estimates were back-transformed into correlation units for reporting and interpretation. Since my new measure captures the full parameter configuration that jointly determines one's ability of response control, this offers a more sensitive test of domain-generalizability.

## **Results**

### **Goal 1: Moving Beyond Isolated Effects of Distractor Strength**

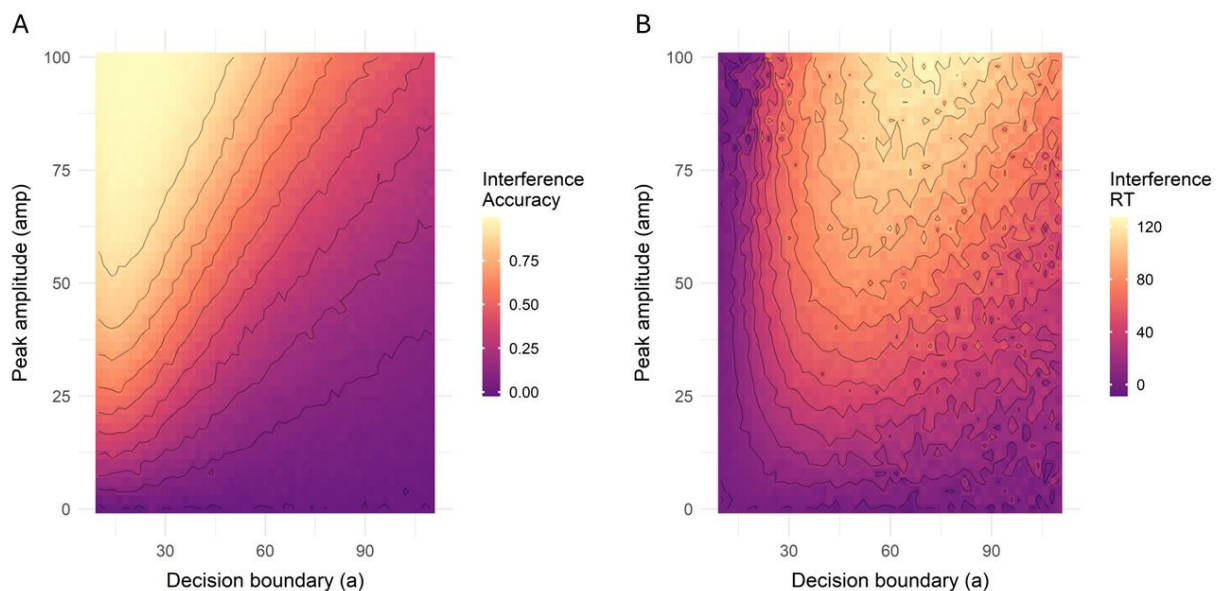
#### ***Amplitude-Boundary Interactions***

To move beyond isolated effects of distractor strength on the degree of behavioral interference one experiences, changes in peak amplitude were investigated alongside variations in decision boundary. As illustrated in Figure 6A, interference effects in accuracy are more pronounced as peak amplitude increases. This is expected given that a larger peak amplitude

indicates larger distractor strength along the automatic route. However, these effects diminish when the height of the decision boundary increases. That the difference in accuracy between congruent and incongruent trials becomes smaller indicates that a higher setting of decision boundary can be beneficial to counteract the influence of task-irrelevant information processed by the automatic route. To formally evaluate this interaction, a linear model was compared with a second-order (quadratic) polynomial regression model predicting accuracy interference effects from peak amplitude, decision boundary and their interaction. Model comparison demonstrated that the quadratic regression model fit the data significantly better than the linear interaction model ( $\Delta\text{BIC} = 4400.17$ ,  $F(5, 2592) = 2338.80$ ,  $p < .001$ ). The quadratic relationship between parameters and interference effects is reflected in Figure 6A by the curvature contour lines. The described heatmap patterns were moreover supported by omnibus tests on the quadratic model, showing a significant main effect of peak amplitude  $F(2, 2592) = 74801.90$ ,  $p < .001$  and decision boundary  $F(2, 2592) = 55752.90$ ,  $p < .001$ , as well as a significant interaction effect between peak amplitude and decision boundary  $F(4, 2592) = 4702.30$ ,  $p < .001$ ) indicating that the effect of distractor strength on accuracy interference effects depends on the setting of decision boundary.

**Figure 6**

*Heatmaps of Simulated Interference Effects in Accuracy and RT across Combinations of Peak Amplitude and Decision Boundary*



*Note.* (A) Interference effects in accuracy, computed as the difference in mean accuracy between congruent and incongruent trials. (B) Interference effects in RT, showing the difference in mean RT between incongruent and congruent trials. Color gradients reflect the

magnitude of each index of behavioral interference as a function of varying levels of peak amplitude and decision boundary. See main text for details.

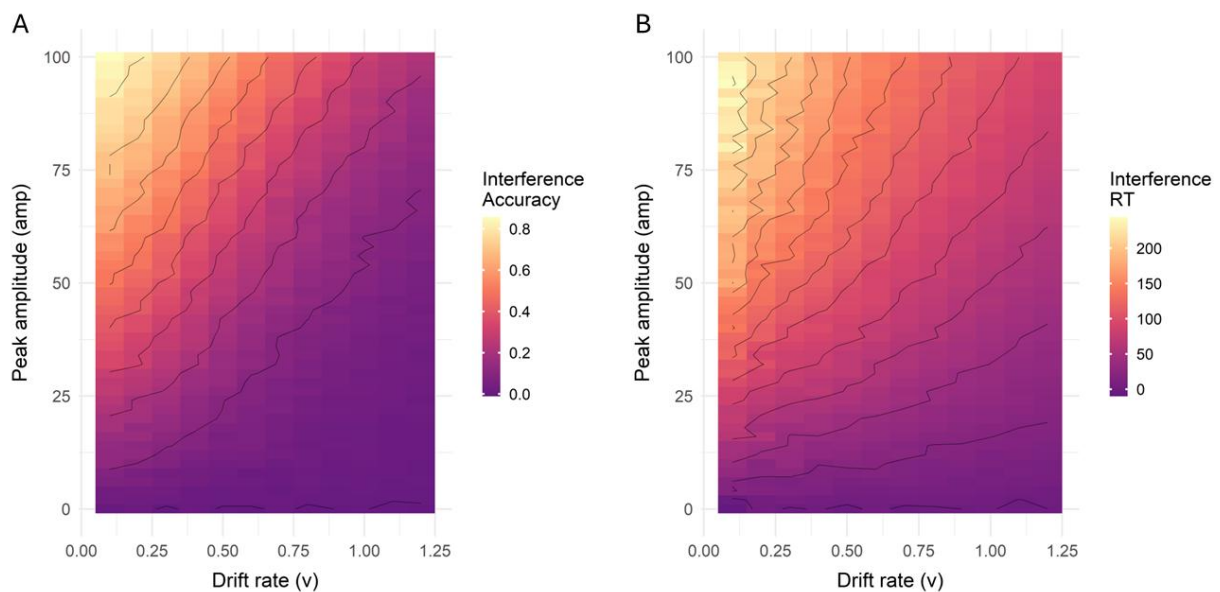
Examining the effects of peak amplitude and decision boundary on behavioral interference costs in RT reveals a complementary picture. When first inspecting the extreme case of a very small peak amplitude in Figure 6B, interference effects in RT change little as decision boundary increases. Intuitively, when the amplitude is near zero, little cognitive conflict is present on incongruent trials. As a result, congruent and incongruent trials are nearly identical, so increasing the amount of evidence to accumulate before deciding affects both trial types in a similar way (resulting in interference effects close to zero). Likewise, at very small decision boundary values, increases in peak amplitude have minimal influence on interference effects in RT. This pattern can be explained by the decision process being finished often before the automatic activation reaches its peak because the decision boundary is set very low. Even large increases in distractor strength thus have limited possibilities to strongly influence the accumulation process, resulting in interference effects near zero. Interestingly, at higher decision boundaries (e.g.,  $a = 60$ ), interference effects grow larger as amplitude increases. In other words, the difference in RT between incongruent and congruent trials becomes more pronounced when task-irrelevant information strongly influences decision formation. However, there seems to be a turning point when decision boundary exceeds peak amplitude. That is, interference effects in RT decrease again when more evidence is accumulated. Put differently, interacting effects between peak amplitude and decision boundary occur. At the same peak amplitude, different levels of interference are observed depending on the height of decision boundary. At the same time, what the optimal setting is of decision boundary to limit the influence of automatic activation depends on how strong peak amplitude is. These patterns are furthermore supported by a quadratic interaction model fitting the data significantly better than a linear interaction model ( $\Delta\text{BIC} = 6088.19$ ,  $F(5, 2592) = 4949.10$ ,  $p < .001$ ). The quadratic model further indicated a significant main effect of peak amplitude  $F(2, 2592) = 22674.60$ ,  $p < .001$  and decision boundary  $F(2, 2592) = 10834.30$ ,  $p < .001$ , as well as a significant interaction effect between peak amplitude and decision boundary  $F(4, 2592) = 2209.40$ ,  $p < .001$ ). This corroborates the above, showing that the effects of peak amplitude on behavioral interference should not be considered without taking the setting of decision boundary into account. In turn, the role of decision boundary as mechanism to manage the net effects of conflicting information cannot be understood as a fixed influence, as its effects on behavioral interference are contingent on the level of distractor strength.

### ***Amplitude-Drift Rate Interactions***

As shown in Figure 7, different levels of interference are observed at the same peak amplitude depending on the magnitude of drift rate. For both accuracy (Figure 7A) and RT (Figure 7B), interference effects are most pronounced when large distractor strength is combined with a low drift rate. Intuitively, when individuals are inefficient at extracting relevant features from a situation of cognitive conflict, distracting information can exert a greater influence on decision formation. However, higher drift rates can counteract the pull of automatic activation toward the incorrect boundary, as reflected in smaller interference effects when drift rate of the controlled route is larger. Furthermore, when someone can extract task-relevant information with much efficiency (i.e., very high drift rate), increasing peak amplitude has minimal distracting effects as indicated by the interference effects near zero.

**Figure 7**

*Heatmaps of Simulated Interference Effects in Accuracy and RT across Combinations of Peak Amplitude and Drift Rate*



*Note.* (A) Interference effects in accuracy, computed as the difference in mean accuracy between congruent and incongruent trials. (B) Interference effects in RT, showing the difference in mean RT between incongruent and congruent trials. Color gradients reflect the magnitude of each index of behavioral interference as a function of varying levels of peak amplitude and drift rate. See main text for details.

Visual inspection of the patterns in Figure 7 is supported by formal analyses of the data. For both accuracy and RT, a second-order polynomial interaction model provided a significantly better fit than a linear interaction model (accuracy:  $\Delta\text{BIC} = 1157.71$ ,  $F(5, 603) = 722.07$ ,  $p < .001$ ; RT:  $\Delta\text{BIC} = 1194.79$ ,  $F(5, 603) = 774.71$ ,  $p < .001$ ). Specifically, significant main effects of peak amplitude (accuracy:  $F(2, 603) = 54886.90$ ,  $p < .001$ ; RT:  $F(2, 603) = 51713.50$ ,  $p < .001$ ) and drift rate (accuracy:  $F(2, 603) = 39868.20$ ,  $p < .001$ ; RT:  $F(2, 603) = 22262.90$ ,  $p < .001$ ) were found. Furthermore, significant interaction effects (accuracy:  $F(4, 603) = 5162.60$ ,  $p < .001$ ; RT:  $F(4, 603) = 2486.10$ ,  $p < .001$ ) support the interpretation that drift rate modulates the extent to which peak amplitude affects behavioral interference and vice versa.

### ***Interim Conclusion***

In sum, these results show that behavioral interference effects in accuracy and RT are shaped by the interacting influence of peak amplitude with decision boundary and drift rate. These findings first highlight the need for computational modeling as a tool to gain deeper insights into response control, given that raw interference effects are not diagnostic about the underlying processes involved in response control. Yet, the present simulations also show that peak amplitude, although often used as a model-based measure of response control, should not be considered in isolation. Even at the same peak amplitude, interference effects can vary substantially depending on the height of decision boundary and magnitude of drift rate. Thus, the effect of peak amplitude is contingent upon the broader parameter configuration of the decision-maker. As the degree of behavioral interference one experiences during cognitive conflict is also shaped by decision strategy and processing efficiency, peak amplitude alone cannot be interpreted in a vacuum in relation to response control. This indicates that in order to understand response control in all its facets, automatic activation should be studied in relation to decision boundary setting and the efficiency of which task-relevant information is processed along the controlled route. Hence, these simulations underscore the need for a novel quantification of response control that captures the broader state of the decision maker (i.e., the interplay of parameters).

### **Goal 2: Toward a Context-Sensitive Measure of Response Control**

The demonstrated interplay of parameters has important implications for common computational measures of response control that adopt a narrow perspective of interpreting peak automatic activation in isolation as a proxy for response control. Specifically, the results of the present study indicate that individuals with similar peak amplitude may nevertheless differ

greatly in how cognitive conflict impacts their decision formation, necessitating a comprehensive approach to quantifying their true control capacity. The following section therefore outlines the development of alternative measures designed to capture how effectively individuals control their responses in the presence of conflicting information. These measures are subsequently evaluated against comparative criteria within this section, with the aim of advancing a context-sensitive conceptualization of response control at the construct level.

### *Development of Candidate Measures*

**Peak-to-Bound Ratio (PTB-Ratio).** To account for the joint influence of peak amplitude and decision boundary, their co-determined impact on decision formation can be captured by the ratio of peak amplitude ( $amp$ ) and decision boundary ( $a$ ):

$$PTB = \frac{amp}{a}$$

The concept of the Peak-To-Bound ratio (PTB-ratio) was originally introduced by Vermeylen et al. (2025), as a measure of sensitivity to task-irrelevant information. The notion of accounting for peak amplitude relative to decision boundary setting is also relevant for the quantification of response control. Specifically, the amount of evidence one accumulates before deciding (i.e., the decision boundary one sets) can override the pull of automatic activation. In other words, as shown in this thesis, the influence of task-irrelevant information depends on how much larger peak amplitude is proportionally to the setting of decision boundary. When this ratio is larger, there is a stronger proportional influence of distractor strength per unit of decision boundary. This indicates a larger net influence of task-irrelevant information on the decision process, reflecting lower response control. In contrast, when PTB-ratio is smaller, the relative influence of distracting information is lower, pointing to a greater ability to control responses during conflict. For example, consider two individuals with identical peak amplitudes of  $amp = 60$ . Person X has an estimated decision boundary parameter of  $a = 90$ , resulting in a ratio of  $PTB = 60/90 = 0.67$ . Person Y, in comparison, has a decision boundary setting at  $a = 40$ , resulting in a ratio of  $PTB = 60/40 = 1.50$ . Although the strength of automatic activation, driven by task-irrelevant information, is identical for both individuals, the higher decision boundary of person X makes the net impact of distractor strength on decision formation smaller. In other words, person X would be identified as showing greater control of responses during cognitive conflict.

**Integrated Evidence Capture (IEC).** To account for the dynamic properties of the unfolding decision process, I developed a novel alternative measure based on the composite evidence accumulation traces. In the DMC, composite evidence traces are based on the superimposed activation of the automatic and controlled route (cf. Figure 3). As the preceding parameter interactions demonstrated, alongside decision boundary, differences in efficiency of evidence accumulation (drift rate; controlled route) can modulate the influence of distractor strength (peak amplitude; automatic route). By constructing a model-based measure of response control based on composite evidence traces, the joint influence of peak amplitude, decision boundary, drift rate and temporal profile is incorporated. To this end, I developed the Integrated Evidence Capture (IEC) as an alternative measure of response control that accounts for the joint interplay of parameters. Formally, the numerator of the IEC corresponds to a mathematical integration, in which the difference between congruent and incongruent evidence traces is summed for every time point  $dt$  (1 ms) until a boundary is reached for the incongruent condition:

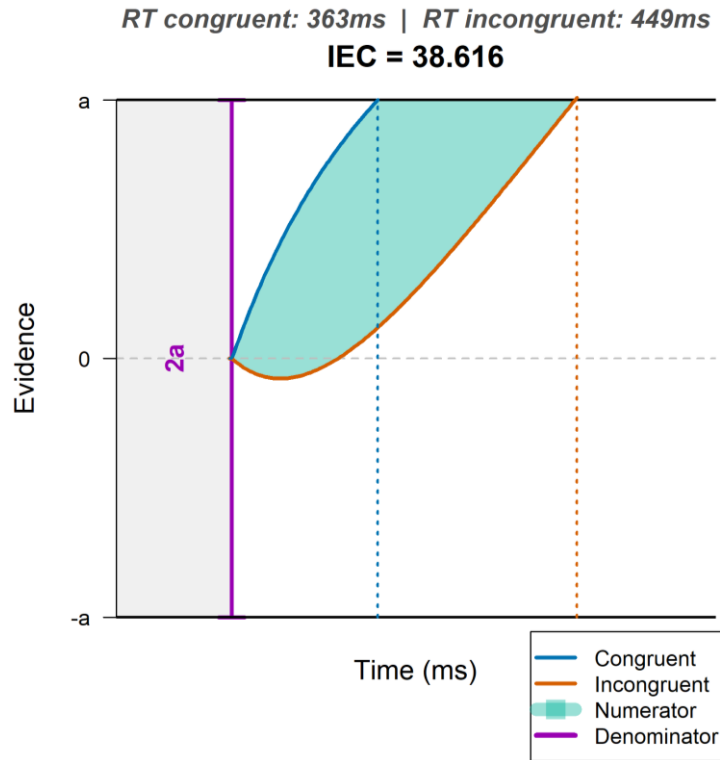
$$IEC = \frac{1}{2a} \int_{t=0}^{t^*} (\min(E_c(t), a) - E_i(t)) dt$$

where  $E_c$  and  $E_i$  denote the congruent and incongruent composite evidence traces at time  $t$  in the decision process, respectively. The congruent evidence trace is capped at the upper decision boundary ( $a$ ), to ensure that the area is constrained to the decision space. By anchoring the integration to the time point of incongruent boundary crossing, this ensures that the measure captures the full history of interference that a decision-maker had to navigate. This is particularly relevant for the dynamics in a Stroop task. Because semantic processing of a color word is slower than spatial processing, cognitive conflict in a Stroop task unfolds later in the decision process. If the integration were cut off at a fixed point, there may be the risk of not capturing the full unfolding of late cognitive conflict (i.e., automatic activation driven by task-irrelevant information). By integrating until the actual realized decision time for the incongruent condition, the IEC accounts for the total area of susceptibility to distracting information that existed up until the very moment a response boundary was reached. This approach ensures that the measure remains sensitive to the influence of distractors, regardless of when in time that influence was most prominent during the decision window. The cumulative difference, representing the area between congruent and incongruent trajectories, is moreover normalized by total boundary separation ( $2a$ ) to evaluate how large the area is between both evidence accumulation traces, relative to the full decision space (see Figure 8). Furthermore,

normalization enables to compare IEC, and thus response control, for different settings of decision boundary.

**Figure 8**

*Integrated Evidence Capture (IEC)*



*Note.* Composite evidence accumulation traces plotted with a boundary setting of  $a = 90$ , drift rate of  $v = 0.8$ , peak amplitude of  $amp = 40$  and time-to-peak parameter of  $\tau = 72$ , mimicking the time course of automatic activation in a Simon task. As a measure of response control at the construct level, IEC quantifies the area between congruent and incongruent evidence trajectories (numerator) relative to the total decision space (denominator).

**Relative Route Regulation (Triple R).** While IEC provides a quantification of response control based on the unfolding dynamics of evidence accumulation, I constructed a third and final candidate measure from a slightly different angle. Specifically, the conceptualization of response control in Triple R is focused on the raw parameter settings of the controlled and automatic route directly. This way, this measure captures how the interacting configuration of parameters of both routes determines one's ability to control responses when navigating cognitive conflict. Put differently, Triple R directly operationalizes response control in terms of the parameters that characterize the decision process, rather than the trajectories that emerge

from it (as in IEC). The findings from the first part of this thesis formed the basis for this third candidate measure. The merit of this approach is that Triple R directly captures the magnitude and directional influence through which DMC parameters jointly shape the net effect of cognitive conflict on decision formation. Moreover, constructing measures from different angles allows for a comparison of these operationalizations, enabling subsequent evaluation of which formulation most adequately characterizes response control at the construct level. Specifically, Triple R quantifies the net influence of distractor strength (i.e., peak amplitude) on decision formation as the combination of two ratios:

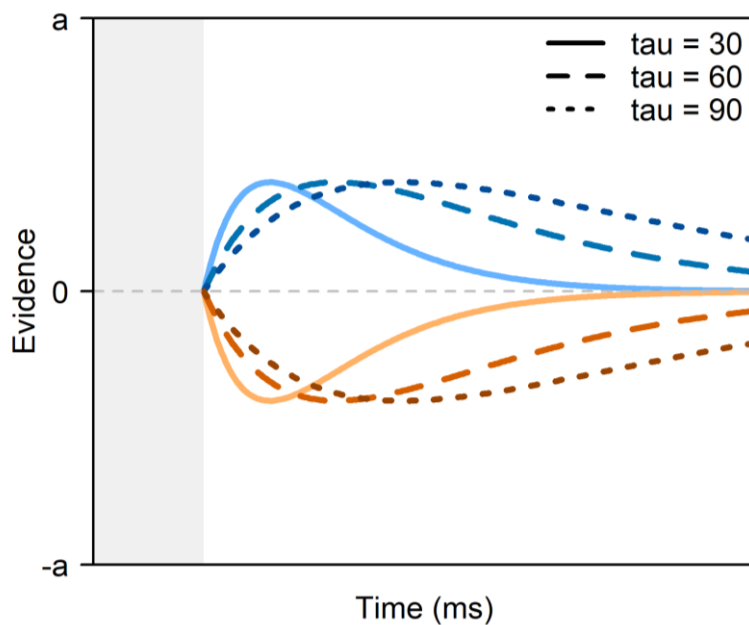
$$Triple\ R = \frac{amp}{a} \cdot \sqrt{\frac{\tau}{(a \cdot v)}}$$

The first ratio  $amp/a$  captures peak amplitude relative to decision boundary, corresponding to the PTB-ratio. Intuitively, it captures the peak of automatic activation, driven by distracting information ( $amp$ ), relative to how much evidence an individual accumulates before committing to a response (i.e., where the decision boundary is set;  $a$ ). As the simulations of amplitude-boundary interactions showed, if distraction peak is nearly equal to or larger than decision boundary, there is a strong pull to the erroneous response option. However, additional evidence accumulation (i.e., higher decision boundary) can counteract this pull, thereby reflecting greater response control. The second ratio was implemented because the simulations of the first part of the current thesis showed that the magnitude of drift rate can modulate the pull of automatic activation toward the incorrect boundary. Specifically, the second ratio  $\sqrt{\tau/(a \cdot v)}$  captures how long automatic activation can influence decision formation ( $\tau$ ), relative to the influence of the controlled route ( $a \cdot v$ ). Crucially, the time-to-peak parameter is situated in the numerator of Triple R, alongside amplitude, because  $\tau$  modulates automatic activation. As illustrated in Figure 9, a larger  $\tau$  means that automatic activation peaks later in the decision-making process. Moreover, automatic activation is more spread out with larger  $\tau$ . As a result, rather than a short-lived influence, the impact of cognitive conflict becomes more persistent and falls within the critical final stages of decision formation. In other words, an individual that has a smaller  $\tau$  estimate would be identified as showing greater control over their responses during cognitive conflict, as automatic activation has only a brief influence on the decision process. This way, the initial pull of distracting information is more effectively counteracted, especially when combined with a larger drift rate and boundary setting. Specifically, by the product of decision boundary ( $a$ ) and drift rate ( $v$ ) in the denominator,

Triple R captures how a higher decision boundary can buy more time for the controlled route to extract task-relevant information more efficiently. Consequentially, a higher  $a$  and/or  $v$  reduces the overall Triple R value, signaling more effective response control because of stronger influence of the controlled route. Additionally, the square root is applied to ensure a linear scaling of time in the second term relative to a similar linear scaling of evidence present in the first term. In sum, because parameters of the automatic route are situated in the numerator and parameters that counteract conflict are placed in the denominator, larger values of Triple R indicate more opportunity for distracting information to interfere with behavior and thus less response control.

**Figure 9**

*The Effect of Tau on Automatic Activation*



*Note.* Visualization of exemplary automatic activation curves generated with a fixed alpha shape parameter ( $\alpha = 2$ ). Larger values of  $\tau$  result in a later peak and more spread-out automatic activation, driven by distracting information. Blue traces denote automatic activation during congruent trials; orange traces during incongruent trials.

### ***Validation of Candidate Measures***

To enable a comparative evaluation of the model-based measures of response control, peak amplitude as stand-alone measure is now evaluated alongside the alternative candidate measures presented above with respect to the following criteria: (1) whether all latent parameters that shape response control are included, (2) whether the index of response control

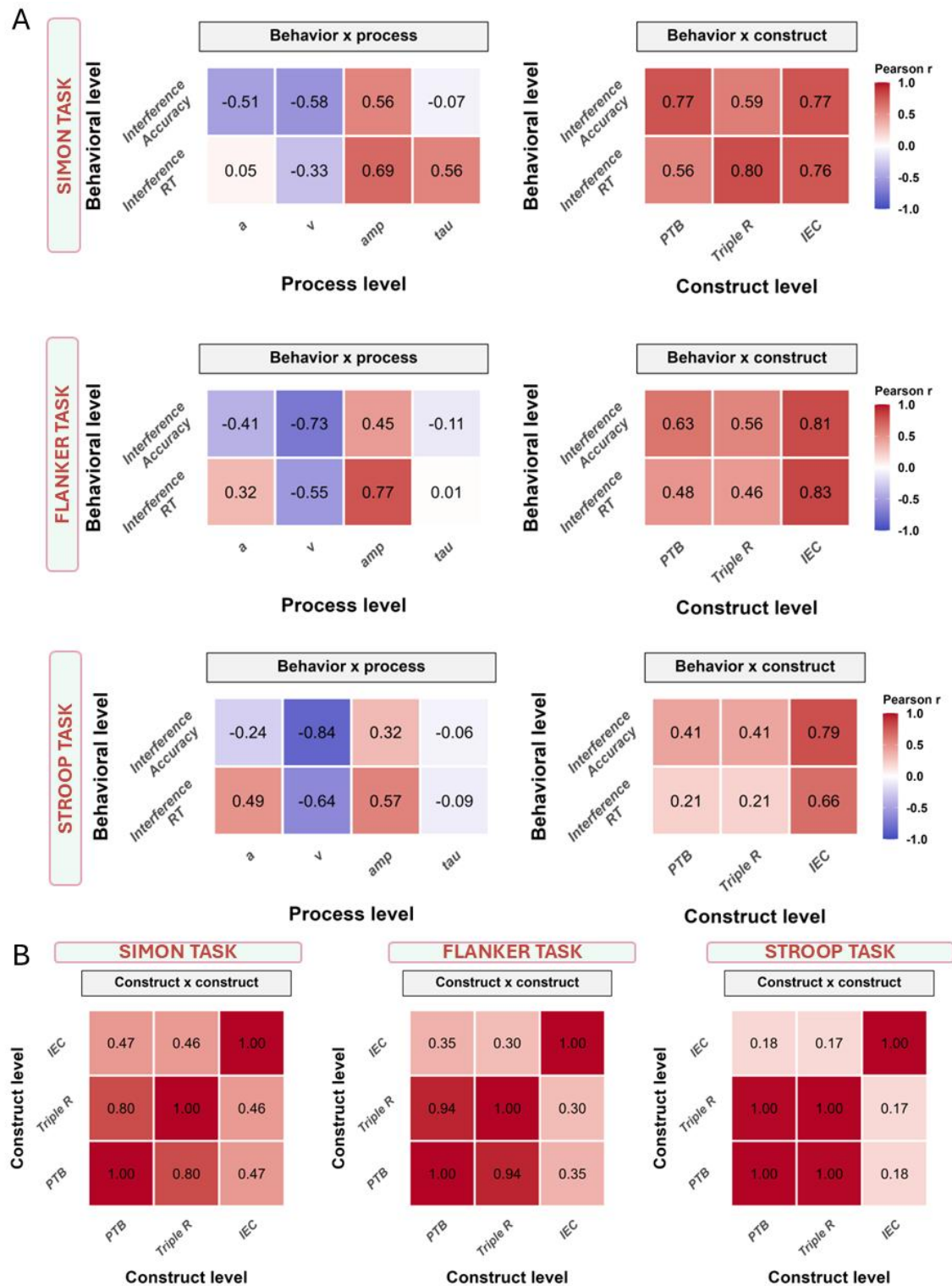
is relevantly tied to behavior and (3) whether it provides a unique mapping between the construct and group differences in task performance, such that observed differences can be explained by the measure of response control.

**Inclusion of All Latent Parameters Shaping Response Control.** When conceptualizing response control as a function of peak amplitude in isolation, the broader decision context is neglected. As demonstrated, the degree of behavioral interference one experiences due to cognitive conflict is jointly determined by peak amplitude, decision boundary and drift rate. In other words, both the amount of evidence one accumulates upon reaching a boundary and the efficiency at which task-relevant evidence accumulation goes are latent processes that interact with the strength of automatic activation to shape response control. Different levels of behavioral interference are linked to the same peak amplitude, depending on the height of decision boundary and magnitude of drift rate (cf. Figure 6-7). These findings indicate that response control cannot be adequately operationalized without considering the broader set of interacting parameters. Against this background, peak amplitude in isolation fails to provide a complete picture of response control, as it does not include all latent parameters that jointly shape one's ability to control responses. Although decision boundary is incorporated as an additional component in the PTB-ratio, differences in efficiency of processing task-relevant information (i.e., drift rate) are not considered. Similarly, the time-to-peak ( $\tau$ ) of automatic activation is not incorporated in the PTB-ratio. Given that automatic activation is more spread-out with larger  $\tau$  (cf. Figure 9), PTB-ratio fails to take into account an important piece of the decision context that modulates the persistence of cognitive conflict. In contrast, IEC and Triple R incorporate the joint influence of peak amplitude, decision boundary, drift rate and time-to-peak. As such, these measures capture the full state of the decision-maker, reflecting the interacting contribution of all latent parameters that shape the degree of response control.

**Meaningful Correlations with Behavioral Interference.** A second criterion for evaluating model-based measures of response control follows from the multi-level conceptual framework of this thesis (cf. Figure 5). Although response control cannot be directly observed in a pure manner (i.e., interference effects capture various factors that are not necessarily specific to the construct of interest), behavioral interference during cognitive conflict can be considered a manifestation of response control as a latent construct. Formally, latent constructs are used to explain behavior (Masyn et al., 2010). This implies that less behavioral interference

should follow from more effective response control, whereas poorer control should be associated with larger interference effects. Therefore, moderate to strong associations between a measure of response control and interference effects are expected. A measure of response control should systematically track behavioral interference, but it should not reduce to interference effects itself. In other words, an adequate measure that captures how well someone can control responses during cognitive conflict should be meaningfully (but not perfectly) tied to behavioral interference.

When first examining peak amplitude (*amp*) in isolation, this stand-alone measure of response control shows moderate to strong positive correlations with interference in accuracy and RT across all three classic conflict tasks (ranging from  $r = 0.32$  to  $r = 0.77$ ; see Figure 10A). This direction of association is expected, given that this DMC parameter indicates the peak strength of distractor-driven automatic activation. Because the alternative measures presented in this thesis are constructed to capture how the joint contribution of interacting parameters shape the degree of response control, it is informative to briefly examine how the other DMC parameters individually relate to behavioral interference. Overall, decision boundary (*a*) shows positive correlations for behavioral interference in RT. Specifically, the correlation is negligible in the Simon task ( $r = 0.05$ ), but moderate in the Flanker ( $r = 0.32$ ) and Stroop task ( $r = 0.49$ ). In contrast, decision boundary shows negative correlations with interference effects in accuracy (Simon:  $r = -0.51$ ; Flanker:  $r = -0.41$ ; Stroop:  $r = -0.24$ ). These patterns indicate that increased response caution reduces errors in the presence of cognitive conflict, but at the cost of longer RTs. Next, moderate to strong negative correlations with both interference effects in accuracy and RT are observed for drift rate (*v*) across all tasks (ranging from  $r = -0.33$  to  $r = -0.84$ ). This means that higher drift rates are associated with a smaller difference in performance between congruent and incongruent trials. Finally, time-to-peak ( $\tau$ ) appears to have limited explanatory value for interference effects as a parameter in isolation, as reflected by the near-zero correlations. Only in the Simon task,  $\tau$  shows a strong positive association with RT interference effects ( $r = 0.56$ ). This suggests that performance in a Simon task is more sensitive to temporal shifts, possibly because automatic activation peaks earlier and is pulse-like. In contrast, in Flanker and Stroop tasks, where automatic activation peaks later and its impact persists across a longer window of the decision process, variations in  $\tau$  appear less critical given that  $\tau$  is already larger in the Flanker and especially in the Stroop task compared to the Simon task.

**Figure 10***Associations Between Behavioral Interference, DMC Parameters and Candidate Measures*

*Note.* A) Pearson correlations between interference effects and DMC parameters (behavior x process level), and between interference effects and alternative model-based measures of response control presented in this thesis (behavior x construct). B) Pearson correlations among the alternative measures of response control (construct x construct). See main text for details.

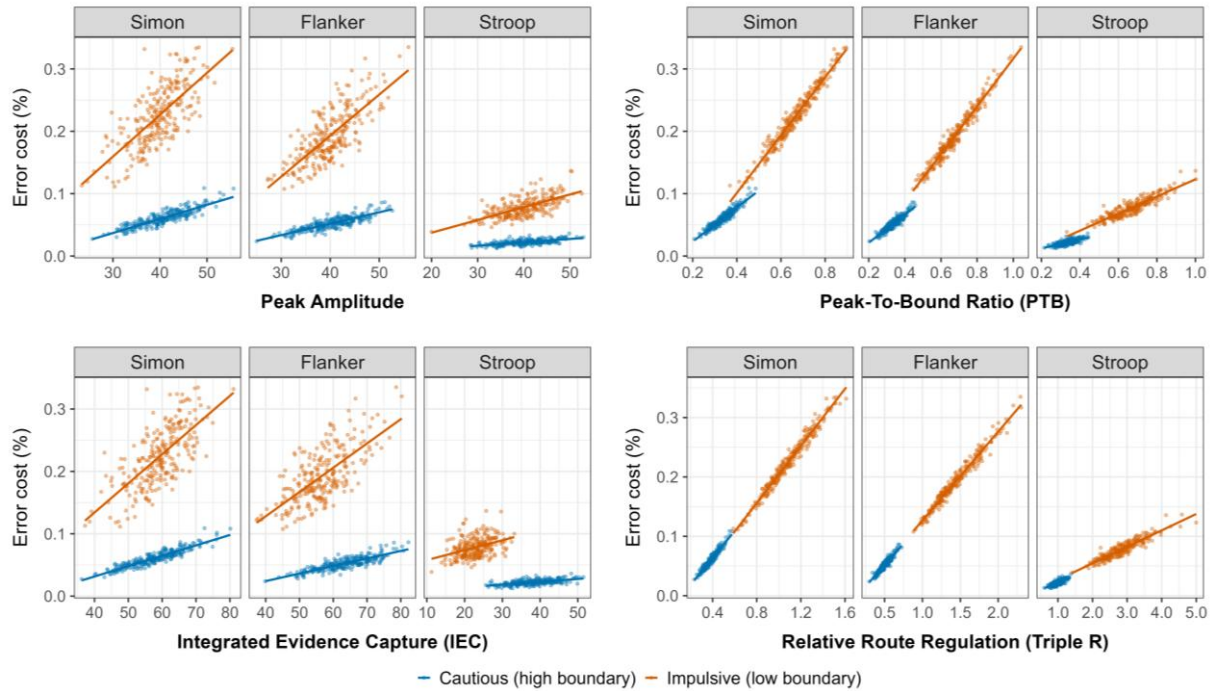
Following the examination of individual DMC parameters, the next step is to evaluate whether the candidate measures presented in this thesis are meaningfully related to behavioral interference. Across the three conflict tasks, PTB-ratio, Triple R and IEC all show positive correlations with behavioral interference in RT and accuracy (see Figure 10A). This indicates that higher values on these measures are associated with larger behavioral interference. This association aligns with their theoretical interpretation that higher values on these measures reflect reduced response control. Examining these correlations more closely, strong correlations are found for all candidate measures in the Simon task, ranging from  $r = 0.56$  to  $r = 0.80$ . The Flanker task shows a similar pattern, with correlations from  $r = 0.46$  to  $r = 0.83$ . In the Stroop task, IEC shows strong correlations with both accuracy and RT interference ( $r = 0.79$  and  $r = 0.66$  respectively), whereas PTB-ratio and Triple R show moderate correlations with accuracy interference ( $r = 0.41$  for both) and weaker correlations with RT interference ( $r = 0.21$  for both). To further examine these patterns, Figure 10B illustrates the associations among these measures at the construct level. This allows to assess the degree of convergence between the candidates and to determine whether they capture unique or shared aspects of behavior. In general, the high overlap between PTB-ratio and Triple R is expected, given that PTB-ratio is a component of Triple R (Simon:  $r = 0.80$ ; Flanker:  $r = 0.94$ ; Stroop:  $r = 1.00$ ). Interestingly, in the Stroop task, PTB-ratio and Triple R capture the same variance in behavioral interference. In contrast, the overlap of IEC with PTB-ratio and Triple R decreases for Flanker and Stroop tasks. Specifically, the overlap between IEC and PTB-ratio decreases from  $r = 0.47$  in the Simon task, to  $r = 0.35$  in the Flanker task, and  $r = 0.18$  in the Stroop task. A similar pattern is found for IEC and Triple R (Simon:  $r = 0.46$ ; Flanker:  $r = 0.30$ ; Stroop:  $r = 0.17$ ), suggesting that IEC captures something distinct in behavior compared to PTB-ratio and Triple R. At the same time, all measures are meaningfully tied to behavior, since the correlations align with their theoretical interpretation as quantifications of response control.

**Unique Mapping to Group Differences.** The final criterion for evaluating the model-based measures of response control is whether they uniquely map onto differences in behavioral performance between groups due to differences in parameter settings that shape response control. For example, if individuals perform worse because they set a lower decision boundary, these individuals should be identified as showing lower control over their responses since the findings of this thesis show that boundary setting is one of the contributing factors that shapes the ability of response control. As shown in Figure 11, peak amplitude as a standalone measure of response control fails to uniquely map onto task performance (i.e. because it is insensitive to

boundary height). For instance, at the same peak amplitude of 40, this measure indicates no difference in response control between the groups with high vs. low decision boundary. Yet, behaviorally, individuals with higher boundaries are less affected by the same level of distractor strength (i.e., peak amplitude) as indicated by a lower error cost. This pattern is confirmed by Receiver Operating Characteristic (ROC) analyses (see Table 2), showing that peak amplitude performs at chance level in discriminating between boundary groups ( $AUC = 0.51-0.54$ ) across tasks. A similar pattern is observed for IEC, where classification performance remains weak in a Simon task ( $AUC = 0.58$ ). Yet, IEC shows improved performance in a Flanker task ( $AUC = 0.75$ ) and almost perfect group separation in a Stroop task ( $AUC \approx 1$ ). This indicates that when automatic activation peaks later, IEC is better able to map the degree of response control onto differences in error cost related to the setting of decision boundary. DeLong tests, comparing AUC values, indeed show that IEC significantly outperforms peak amplitude ( $z = 19.23, p < .001$ ). Furthermore, near-perfect classification across all tasks is found for PTB-ratio and Triple R ( $AUC \approx 1$ ). This indicates that both measures are able to almost perfectly discriminate between high vs. low boundary groups.

**Figure 11**

*Mapping of Response Control Measures and Error Cost for Group Differences in Decision Boundary*



*Note.* For a group of cautious decision-makers (i.e., high decision boundary), data is simulated with a group mean of  $a = 120$ , while  $a = 60$  for a group of individuals with low boundary

setting. Drift rate and peak amplitude for both groups are sampled from the same normal distributions, with the aim of no systematic group differences outside of decision boundary ( $v = 0.5$ ;  $amp = 40$ ). Only PTB-ratio and Triple R ensure a one-to-one mapping, as values on these measures of response control do not overlap between groups of high vs. low decision boundary that differ in task performance (i.e., error cost).

**Table 2**

*ROC-Based Classification of High vs. Low Decision Boundary and DeLong Comparisons of Response Control Measures Relative to Peak Amplitude*

|                     | AUC  | <i>z</i> | <i>p</i> |
|---------------------|------|----------|----------|
| <b>Simon task</b>   |      |          |          |
| Peak amplitude      | 0.54 |          |          |
| PTB-ratio           | 1.00 | 17.49    | < .001   |
| IEC                 | 0.58 | 0.66     | .5076    |
| Triple R            | 1.00 | 17.73    | < .001   |
| <b>Flanker task</b> |      |          |          |
| Peak amplitude      | 0.54 |          |          |
| PTB-ratio           | 1.00 | 17.94    | < .001   |
| IEC                 | 0.75 | 20.82    | < .001   |
| Triple R            | 1.00 | 17.94    | < .001   |
| <b>Stroop task</b>  |      |          |          |
| Peak amplitude      | 0.51 |          |          |
| PTB-ratio           | 1.00 | 18.54    | < .001   |
| IEC                 | 1.00 | 19.23    | < .001   |
| Triple R            | 1.00 | 18.86    | < .001   |

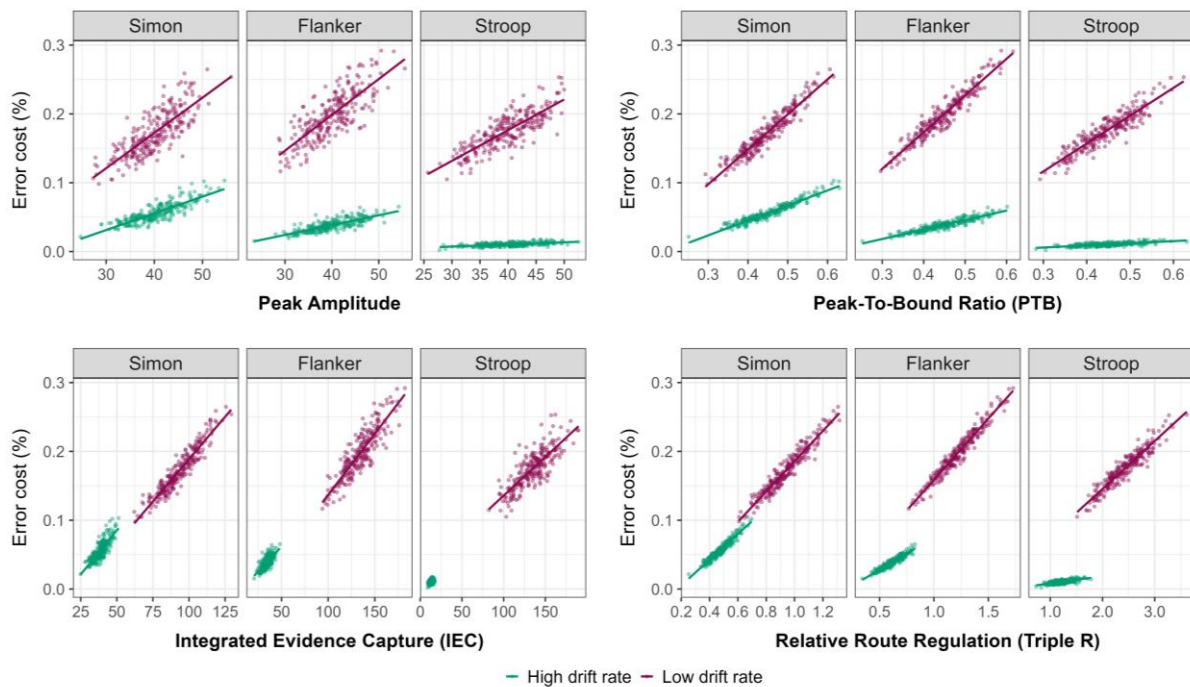
*Note.* AUC = area under the Receiver Operating Characteristic (ROC) curve. Higher AUC values indicate a better classification of boundary groups (high vs. low decision boundary). DeLong tests compare each measure against peak amplitude. Positive *z*-values indicate that the alternative response control measure shows better classification performance than peak amplitude.

Next to differences in decision boundary, studies have also reported differences between groups in processing of task-relevant information (e.g., Patel et al., 2025). Furthermore, the findings of this thesis show that the magnitude of drift rate along the controlled route can

counteract the degree of behavioral interference one experiences. Therefore, the same simulation approach as for decision boundary was adopted for two groups with high vs. low drift rate. As Figure 12 shows, peak amplitude assigns overlapping values of response control to both groups despite differences in performance related to drift rate. This is reflected in the ROC analyses, which show classification performance around chance level for peak amplitude ( $AUC = 0.50-0.53$ ) across the classic conflict tasks. This suggests that peak amplitude as a measure of response control cannot explain performance variation related to the specific component (i.e., drift rate) that shapes response control. Similarly, PTB-ratio fails to provide a one-to-one mapping, with AUC values around chance level ( $AUC = 0.50-0.52$ ). Only IEC and Triple R successfully distinguish between both groups across tasks ( $AUC \approx 1$ ). Specifically, the group with low drift rate experiences more behavioral interference, as reflected in higher error costs (Figure 12). This group is assigned a higher value on both IEC and Triple R, indicating poorer response control. The response control values assigned to the low drift rate group do not overlap with those of high drift rate, showing a clear separation. As shown in Table 3, IEC and Triple R demonstrate significantly higher classification performance relative to peak amplitude.

**Figure 12**

*Mapping of Response Control Measures and Error Cost for Group Differences in Drift Rate*



*Note.* For a group with greater processing efficiency of task-relevant information (i.e., high drift rate), data is simulated with a group mean of  $v = 0.8$ , while  $v = 0.2$  for a group with lower

processing efficiency (i.e., low drift rate). Decision boundary and peak amplitude for both groups are sampled from the same normal distributions, with the aim of no systematic group differences outside of drift rate ( $a = 90$ ;  $amp = 40$ ). Only for IEC and Triple R, a one-to-one mapping exists between response control as quantified by these measures and group differences in error cost related to drift rate.

**Table 3**

*ROC-Based Classification of High vs. Low Drift Rate and DeLong Comparisons of Response Control Measures Relative to Peak Amplitude*

|                     | AUC  | <i>z</i> | <i>p</i> |
|---------------------|------|----------|----------|
| <b>Simon task</b>   |      |          |          |
| Peak amplitude      | 0.52 |          |          |
| PTB-ratio           | 0.52 | -0.13    | 0.8959   |
| IEC                 | 1.00 | 18.61    | < .001   |
| Triple R            | 1.00 | 18.52    | < .001   |
| <b>Flanker task</b> |      |          |          |
| Peak amplitude      | 0.50 |          |          |
| PTB-ratio           | 0.50 | 0.15     | 0.8771   |
| IEC                 | 1.00 | 19.29    | < .001   |
| Triple R            | 1.00 | 19.30    | < .001   |
| <b>Stroop task</b>  |      |          |          |
| Peak amplitude      | 0.53 |          |          |
| PTB-ratio           | 0.51 | -1.46    | 0.1430   |
| IEC                 | 1.00 | 18.12    | < .001   |
| Triple R            | 1.00 | 18.08    | < .001   |

*Note.* AUC = area under the Receiver Operating Characteristic (ROC) curve. Higher AUC values indicate a better classification of groups that differ in the efficiency of task-relevant evidence accumulation (high vs. low drift rate). DeLong tests compare each measure against peak amplitude. Positive *z*-values indicate that the alternative response control measure shows better classification performance than peak amplitude.

### ***Interim Conclusion***

The evaluation of the different measures of response control across the three criteria is summarized in Table 4. First, only IEC and Triple R incorporate the joint influence of the peak

amplitude, decision boundary, drift rate and time-to-peak, thereby taking into account the full state of the decision-maker. Second, all measures show meaningful associations with behavioral interference in the expected direction, supporting their interpretation as quantifications of response control. Third, only Triple R is sensitive to explain differences in task performance related to specific components (i.e., decision boundary and drift rate) that shape response control. Altogether, the evaluation criteria point to Triple R as the most robust, novel quantification of response control at the construct level.

**Table 4**

*Overview of Response Control Measures and Evaluation Criteria*

|                | Inclusion of all latent<br>parameters | Meaningful behavioral<br>correlations | Unique mapping to<br>group differences |
|----------------|---------------------------------------|---------------------------------------|----------------------------------------|
| Peak amplitude | ×                                     | ✓                                     | ×                                      |
| PTB-ratio      | ×                                     | ✓                                     | ×                                      |
| IEC            | ✓                                     | ✓                                     | ×                                      |
| Triple R       | ✓                                     | ✓                                     | ✓                                      |

### Goal 3: Revisiting Domain-Generality of Response Control

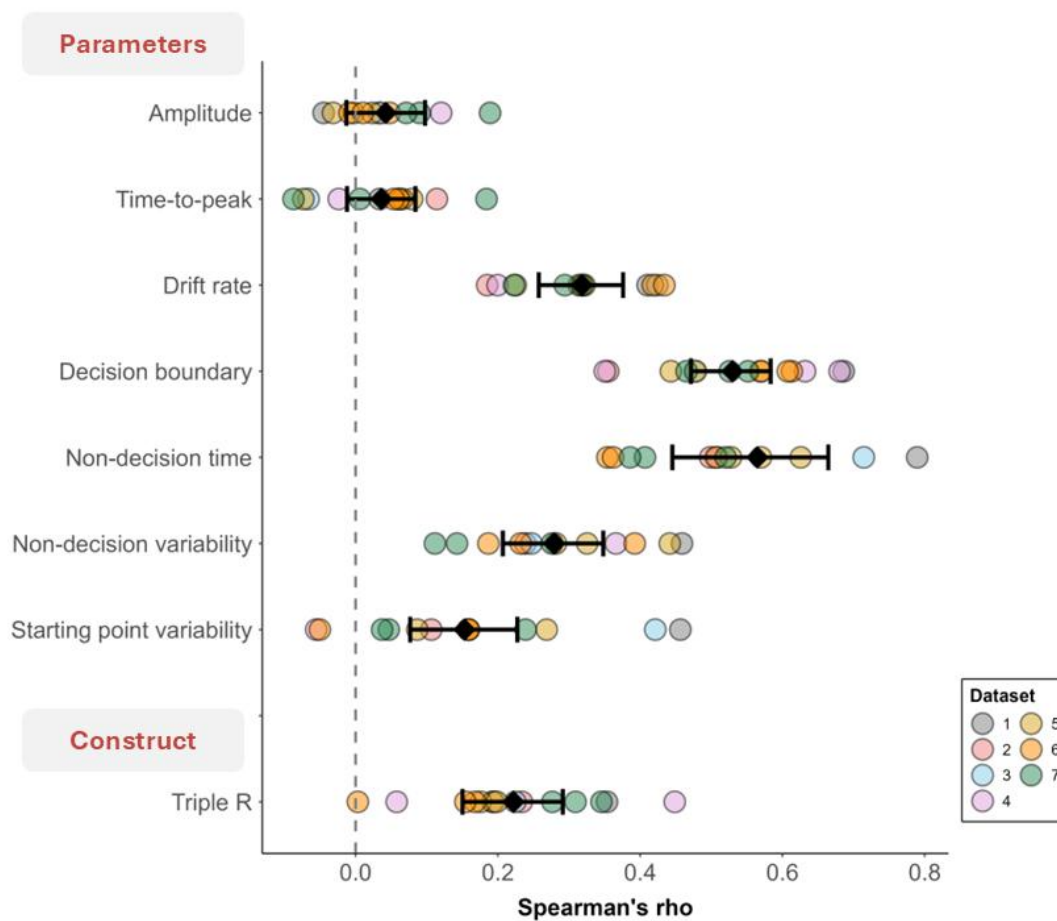
Whereas a previous influential study by Hedge et al. (2022) questioned the domain-general nature of response control, these conclusions were based on (the absence of between-task correlations of) the parameters of automatic activation in isolation. However, the findings of this thesis highlight the importance of taking the full state of the decision maker (i.e., the interplay of parameters) into account to provide a full picture of response control. Therefore, I revisited the question of domain-generality, with response control formalized as the net impact of distractor strength on decision formation as captured by my novel measure Triple R.

Similar as reported by Hedge et al. (2022), the results of the multilevel random effects meta-analysis (see Figure 13) indicate a very small and non-significant positive correlation for peak amplitude ( $r = 0.04$ , 95% CI =  $[-0.01, 0.10]$ ,  $p = 0.13$ ) and time-to-peak parameter ( $r = 0.04$ , 95% CI =  $[-0.01, 0.08]$ ,  $p = 0.14$ ). Moderate to strong positive correlations across conflict tasks are found for drift rate ( $r = 0.32$ , 95% CI =  $[0.26, 0.38]$ ,  $p < .001$ ) and decision boundary ( $r = 0.53$ , 95% CI =  $[0.47, 0.58]$ ,  $p < .001$ ). Next, significant positive correlations can be observed for non-decision time ( $r = 0.56$ , 95% CI =  $[0.45, 0.66]$ ,  $p < .001$ ), variability of non-decision time ( $r = 0.28$ , 95% CI =  $[0.21, 0.35]$ ,  $p < .001$ ) and starting

point variability ( $r = 0.15$ , 95% CI =  $[0.08, 0.23]$ ,  $p < .001$ ). Together, these parameter-level results replicate the findings reported by Hedge et al. (2022). The critical question, however, is whether response control shows evidence of domain-generalty when examined at the construct level rather than at the level of isolated model parameters. Importantly, when response control is quantified using the novel Triple R measure, which captures the full decision context in which cognitive conflict occurs, a significant positive correlation is observed ( $r = 0.22$ , 95% CI =  $[0.15, 0.29]$ ,  $p < .001$ ). This indicates that individuals who display greater response control in one conflict task (as measured by the net impact of distractor strength on decision formation), will also be more likely to exhibit greater response control in other conflict tasks.

**Figure 13**

*Forest Plot of Correlations Between Task Pairs*



*Note.* Error bars represent 95% confidence intervals. Colored circles indicate individual correlations between task pairs within each dataset, while black diamonds show the overall meta-analytic correlation across datasets.

## **Discussion**

### **Distractor Strength is Only Part of the Story**

The present work identified and addressed a conceptual problem in current literature of cognitive conflict: the notion of response control is predominantly approached through indices of automatic activation in isolation. To move beyond this isolated perspective, the first section of this thesis studied the effects of distractor strength on the degree of behavioral interference within the broader decision context it occurs. Through a simulation study with the Diffusion Model for Conflict tasks (DMC; Ulrich et al., 2015), regression analyses revealed significant interactions between decision boundary and peak amplitude (i.e., distractor strength) on interference effects in both accuracy and RT. In line with the visual analysis of the heatmaps, a similar pattern emerged for different combinations of peak amplitude and drift rate. Dependent on the strength of distracting information and the efficiency of evidence accumulation, different levels of interference effects (in accuracy and RT) were observed, as supported by significant interaction effects between peak amplitude and drift rate. Altogether, since the height of the decision boundary and magnitude of drift rate relative to peak amplitude can restrain the impact of distracting information on decision formation, these findings underscore the importance to account for the interplay of parameters to provide a complete picture of response control. This has important implications for common computational approaches used to measure response control that focus on isolated parameters of the automatic route (e.g., Hedge et al., 2022; Kelber et al., 2025; Luo et al., 2024; Servant & Evans, 2020). In particular, these simulations show that response control cannot be directly quantified by simply looking at the strength of automatic activation, since its net impact depends on its interaction with other parameters. As a result, neglecting the broader contribution of parameters may lead to incomplete inferences about response control. Specifically, individuals with similar peak amplitudes may differ greatly in the extent to which they can control their responses during cognitive conflict. This shows the need for a novel operationalization of response control that takes the full state of the decision-maker (i.e., the interplay of parameters) into account.

### **Measuring Response Control at the Construct Level**

Building on the parameter interactions, the second part of the current work focused on the development and evaluation of alternative model-based measures of response control to capture this control ability at the construct level, rather than through isolated parameters. Three candidate measures of response control were developed, each constructed from a slightly different angle, to provide a more context-sensitive quantification of response control. The

comparative evaluation of these measures, next to peak amplitude, revealed several key insights.

### ***Peak Amplitude as an Incomplete Measure of Response Control***

The simulation-based analyses to evaluate the measures demonstrated a fundamental limitation of inferring response control from peak automatic activation in isolation. Since peak amplitude as stand-alone measure neglects the interacting effects of other parameters, it fails to explain performance variation related to the specific components (i.e., decision boundary and drift rate) that shape response control. This lack of one-to-one mapping was consistently found across Simon, Flanker and Simon tasks. Second, the importance of integrating the full state of the decision-maker is further illustrated by the correlations between parameters and interference effects obtained from DMC-based simulations. Specifically, decision boundary was negatively correlated with interference effects in accuracy. That is, the difference in accuracy between congruent and incongruent trials tends to be smaller when more evidence is accumulated. This further shows the role of boundary setting that can act as a controlling mechanism to reduce the impact of conflicting information during decision formation. Notably, the strength of this negative correlation showed a gradual attenuation across conflict tasks, from Simon to Flanker to Stroop. This suggests that the additional benefit of a higher decision boundary may become smaller as the peak of automatic activation occurs later and is more spread out. In other words, a higher setting of decision boundary generally reduces interference effects in accuracy, but the effect in itself appears less pronounced in tasks with later and more sustained automatic activation. This highlights the importance of accounting for the specific task context in which cognitive conflict occurs. Next to decision boundary, the role of drift rate in response control was moreover reflected in its negative correlations with interference effects across all classic conflict tasks. Consistent with the reported amplitude-drift rate interactions, this shows that increased efficiency of evidence accumulation can counteract the influence of task-irrelevant information, as higher drift rates were associated with a smaller difference in performance between congruent and incongruent trials.

### ***Accounting for All Parameters is Necessary but Not Sufficient***

The comparative evaluation of the model-based measures further showed that incorporating all latent parameters that interactively shape the degree of response control is a necessary but not a sufficient prerequisite for a valid measure of the construct. Specifically, compared to peak amplitude, Peak-To-Bound ratio (PTB-ratio) provides a more nuanced conceptualization of response control by explicitly accounting for the role of the decision

boundary. While no unique mapping exists between peak amplitude and behavioral performance, PTB-ratio successfully explained variations in task performance that were related to differences in the setting of decision boundary. However, this measure assigned similar values of response control to two groups with high vs. low drift rate, despite differences in performance specifically related to differences in drift rate. This shows that PTB-ratio, originally introduced by Vermeulen et al. (2025) to measure one's sensitivity to task-irrelevant information, cannot be directly adopted as a measure of response control. As PTB-ratio does not incorporate the influence of drift rate and task-specific time course of automatic activation, this highlights the need to take the full state of the decision-maker (i.e., all latent parameters) into account. Accordingly, the novel Integrated Evidence Capture (IEC) measure integrates the influence of the full set of parameters by design. Because it is based on composite evidence accumulation trajectories, it should theoretically offer a complete picture of response control. However, IEC was not able to distinguish simulated groups of high vs. low boundary setting across all classic conflict tasks. Put differently, this indicates that IEC fails to consistently explain group differences in task performance driven by boundary variations, even though decision boundary shapes the influence of distracting information on decision formation.

### ***Triple R as a Novel Context-Sensitive Measure of Response Control***

The evaluation of candidate measures revealed that meaningful associations with behavioral interference are a necessary, but not a sufficient prerequisite for a valid measure of response control. Although all quantifications of response control showed moderate to strong correlations with interference effects in the expected direction, only the novel Triple R measure consistently showed near-perfect classification performance for different groups, across all conflict tasks. Specifically, across Simon, Flanker and Stroop tasks, Triple R successfully distinguished simulated groups that differed in decision boundary and drift rate by assigning distinct levels of response control, in line with the group differences in behavioral interference. Importantly, these mappings were non-overlapping, indicating that variations in task performance could be uniquely explained by differences in the interacting configuration of latent parameters captured by Triple R.

Altogether, the simulation-based results demonstrate the importance of taking the full state of the decision-maker (i.e., the interplay of parameters) into account to capture response control in all its facets. Importantly, these findings indicate that even if a measure includes all latent parameters that jointly shape the net impact of distracting information on decision formation, this does not automatically guarantee an adequate quantification of response control.

Similarly, next to the importance of being meaningfully associated with behavioral outcomes, a valid measure of response control should appropriately map differences in latent response control ability onto differences in task performance that arise from specific parameter configurations shaping the decision process. In this regard, Triple R emerged as the most robust quantification of response control, meeting all three evaluation criteria outlined in the current work. By combining the influence of peak amplitude and time-to-peak of automatic activation in the numerator, with compensatory effects of decision boundary and drift rate in the denominator, Triple R directly formalizes how well someone can control their responses during cognitive conflict.

### **Implications for Domain-Generality**

The broader implications of the current findings are illustrated in the third and final section of the present work, on the on-going debate whether a common response control mechanism exists. Previous work by Hedge et al. (2022) questioned the domain-general nature of response control based on non-significant correlations for peak amplitude and time-to-peak of automatic activation across conflict tasks. Since they used an isolated parameter approach, this question was revisited in light of the present thesis which underscores the importance of accounting for the interplay of parameters. The present results shed important new light on the on-going debate regarding the domain-general nature of response control. Specifically, when response control is conceptualized as the net impact of distractor strength on decision formation as captured by Triple R, a significant positive correlation between conflict tasks was found. In other words, when the full state of the decision-maker is taken into account, and thus all parameters that jointly shape response control are incorporated, a different picture emerges. These findings indicate that individuals who display greater response control in one conflict task are also more likely to exhibit greater response control in other conflict tasks. In other words, these results provide evidence for the domain-general nature of response control. This connects directly to broader theories of executive functioning. Specifically, in the original three-factor model proposed by Miyake et al. (2000), executive functions are characterized by unity and diversity. Inhibition is one of the three factors in this model, reflecting shared variance in performance across tasks. Although the current analysis focused on classic conflict tasks rather than the broader set of inhibition tasks, such as the anti-saccade task (Hallett, 1978) or stop-signal task (Logan, 1985), the significant positive correlation of Triple R between conflict tasks is consistent with the broader notion of a common mechanism to control responses in different contexts. In this regard, this suggests that previous conclusions against domain-generality have

been (partly) dependent on incomplete operationalizations of response control. The absence of correlations across tasks has moreover led to the debate whether inhibition as a latent construct can even be reliably estimated (Draheim et al., 2021; Hedge et al., 2022; Rey-Mermet et al., 2018). However, the present work shows that when response control is captured in all its facets, as measured by the novel Triple R measure, evidence does emerge for a common response control mechanism across tasks. This suggests that previous evidence against domain-generalizability of response control should not necessarily be interpreted as evidence against the value of response control as a psychometric construct but may instead reflect limitations of the measures used to capture it. In sum, this empirical application demonstrates the value of Triple R as a newly developed measure of response control that can be used to address important fundamental questions.

### **Limitations and Future Directions**

Despite the robust profile of Triple R across the evaluation criteria outlined in this work, a notable observation in the simulation-based results of the second criterion was the attenuated correlations of Triple R (and PTB-ratio) with interference effects in the Stroop task. Moreover, Triple R and PTB-ratio captured the exact same variance of behavioral interference in this task-context. Given that PTB-ratio is a component of Triple R, this suggests that the second ratio in Triple R, that accounts for the temporal dynamics of automatic activation relative to the controlled route, did not explain any additional variance of behavior. These patterns may be attributed to a limitation in the parameter settings used for these simulations. Although the mean and standard deviations of the distributions from which the parameters were sampled align with those reported by Hedge et al. (2022), these distributions were the same to simulate data for Simon, Flanker and Stroop Tasks. Task-specific differences were only limited to the differences in the mean of the time-to-peak parameter that modulates automatic activation. Consequentially, the simulations assumed that an individual approaches a Stroop trial with the same level of caution (i.e., decision boundary) and efficiency of task-relevant information (i.e., drift rate) as they do during a Simon or Flanker trial. As specified in the methods, data was simulated using a low peak amplitude ( $\mu = 27.5$ ) and a relatively high drift rate ( $\mu = 0.5$ ). These parameter settings in combination with a mean decision boundary of 62.5 result in fast decisions, with an average decision time of 125ms. When this fast decision-making is paired with a late peak automatic activation as in a Stroop task ( $\mu = 505$ ), it is likely that the decision process already ended before cognitive conflict could fully unfold. This may explain why the second ratio in Triple R, intended to capture these task-specific temporal dynamics, became

redundant and captured the exact same variance as the PTB-ratio. As a result, the attenuated correlations of Triple R with behavioral interference for the Stroop task might be a byproduct of these specific parameter settings. Specifically, in the DMC fits reported by Hedge et al. (2022), empirical Stroop drift-rate means were closer to 0.28-0.33. Thus, the simulated Stroop data, with a mean drift rate of 0.5, is based on more efficient evidence accumulation. This is important because a higher drift rate shortens decision time. At the same time, these considerations further underscore the importance of considering the context and thus the full set of interacting (task-specific) parameters in order to study response control in all its facets. Therefore, a useful next step is to simulate data for different settings of more task-specific parameters to assess the conditions in which the second component of Triple R explains additional variance in behavior, beyond PTB-ratio in a Stroop task.

When examining Triple R in Figure 11-12, it is noteworthy that Triple R values are consistently higher for the Stroop task compared to the Flanker and especially the Simon task. Looking at the formula of this measure, the time-to-peak parameter is situated in the numerator. Since this parameter is larger for a Stroop task than for a Flanker or Simon task, the construction of Triple R is based on the assumption that response control is generally lower in a task in which automatic activation unfolds later and is more sustained over time. This comes with a critical note that it remains unclear whether this is an artefact of how Triple R incorporates the time-to-peak parameter or whether this reflects a true difference in task-specific demands on response control. Yet, the consistently higher Triple R values observed for the Stroop task compared to the Flanker and Simon task can be understood in light of the dimensional-overlap theory (Kornblum, 1992; Kornblum et al., 1990). In this taxonomy, the Stroop task is classified as a task context in which there is overlap between relevant and irrelevant stimulus dimensions (i.e., ink color and word meaning) themselves and between stimulus-response dimensions. In contrast, the Simon task is only characterized by overlapping stimulus-response dimensions. In other words, there are more sources of conflict within the Stroop task, which may together be more complex to control. Hence, this raises an interesting question for future research to investigate the extent to which response control is more demanding in a Stroop task due to its late, semantic nature of cognitive conflict.

A key feature of the newly developed Triple R measure is that it consistently provided a unique mapping of latent response control ability onto observed performance differences between groups across the classic conflict tasks. In other words, Triple R is able to distinguish individuals with greater vs. weaker response control because of differences in the settings of latent parameters that shape performance during cognitive conflict. This distinction is

particularly relevant in clinical contexts, since impairments in response control play a key role in multiple psychiatric disorders such as obsessive-compulsive spectrum disorders (van Velzen et al., 2014) and depression (Gohier et al., 2009). In contrast to previous approaches that quantify response control based on interference effects or indices of automatic activation in isolation, Triple R takes the full state of the decision-maker into account. Specifically, by construction it incorporates the joint influence of peak amplitude, decision boundary, drift rate and time-to-peak parameter. As a result, Triple R is not only able to identify impairments in response control but also has the potential to identify why such difficulties emerge. Therefore, in future research, I plan to work toward a diagnostic framework that specifies how differences in control ability can be traced back to parameter-specific roles played by distractor strength, processing efficiency, decision strategy and time course of automatic activation. This way, Triple R would not only be able to identify who displays greater response control, but also to provide nuanced insights into what exactly makes someone better to control the net influence of irrelevant information on decision formation.

## **Conclusion**

To conclude, the findings of the current work advance our understanding and measurement of response control at the construct level. Specifically, these insights show that response control in all its facets cannot be understood through indices of automatic activation alone but is shaped by the dynamic interplay of latent parameters that jointly influence how distracting information impacts decision formation. The present work proposes Triple R as a novel, context-sensitive measure of response control that takes the full state of the decision-maker into account. This moves the field to approach response control from an integrative perspective and offers nuanced insights to work toward a diagnostic framework relevant for clinical practice.

## References

- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290–303. <https://doi.org/10.1037/0882-7974.19.2.290>
- De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 731–750. <https://doi.org/10.1037/0096-1523.20.4.731>
- Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General, 150*(2), 242–275. <https://doi.org/10.1037/xge0000783>
- Ellinghaus, R., Liepelt, R., Mackenzie, I. G., & Mittelstädt, V. (2024). Conflict processing is unaffected by stimulus duration across multiple visual tasks: Evidence for transient over permanent activation models. *Computational Brain & Behavior, 7*(3), 378–398. <https://doi.org/10.1007/s42113-024-00211-x>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Farell, S., & Lewandowsky, S. (Eds.). (2015). *An introduction to cognitive modeling*. In B. U. Forstmann, & E.-J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 3-24). Springer. <https://doi.org/10.1007/978-1-4939-2236-9>
- Fröber, K., & Lerche, V. (2023). Performance-contingent reward increases the use of congruent distracting information. *Attention, Perception, & Psychophysics, 85*(3), 905–929. <https://doi.org/10.3758/s13414-023-02682-9>

- Gerlach, C., & Poirel, N. (2026). Do congruency effects in Navon's and Stroop's paradigms reflect the same (control) mechanisms? A conceptual analysis and an empirical test. *Psychological Review*. <https://doi.org/10.1037/rev0000617>
- Gohier, B., Ferracci, L., Surguladze, S. A., Lawrence, E., El Hage, W., Kefi, M. Z., Allain, P., Garre, J.-B., & Le Gall, D. (2009). Cognitive inhibition and working memory in unipolar depression. *Journal of Affective Disorders*, 116(1), 100–105. <https://doi.org/10.1016/j.jad.2008.10.028>
- Haller, S. P., Stoddard, J., Pagliaccio, D., Bui, H., MacGillivray, C., Jones, M., & Brotman, M. A. (2021). Computational modeling of attentional impairments in disruptive mood dysregulation and attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 60(5), 637–645. <https://doi.org/10.1016/j.jaac.2020.08.468>
- Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, 18(10), 1279–1296. [https://doi.org/10.1016/0042-6989\(78\)90218-3](https://doi.org/10.1016/0042-6989(78)90218-3)
- Hedge, C., Powell, G., Bompas, A., & Sumner, P. (2022). Strategy and processing speed eclipse individual differences in control ability in conflict tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 48(10), 1448–1469. <https://doi.org/10.1037/xlm0001028>
- Hofmann, M. A. (2016). Null hypothesis significance testing in simulation. *2016 Winter Simulation Conference (WSC)*, 522–533. <https://doi.org/10.1109/WSC.2016.7822118>
- Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological Review*, 117(3), 759–784. <https://doi.org/10.1037/a0019471>
- Janczyk, M., Mackenzie, I. G., Ulrich, R., & Koob, V. (2025). *Ten years diffusion model for conflict (DMC) tasks: Theoretical foundations, applications, practical*

*recommendations, and open challenges.* PsyArXiv.

[https://doi.org/10.31234/osf.io/rgepd\\_v1](https://doi.org/10.31234/osf.io/rgepd_v1)

- Kelber, P., Mittelstädt, V., & Ulrich, R. (2025). Interplay of aging and practice in conflict processing: A big-data diffusion-model analysis. *Psychology and Aging*, 40(1), 66–85. <https://doi.org/10.1037/pag0000848>
- Khng, K. H., & Lee, K. (2014). The relationship between Stroop and stop-signal measures of inhibition in adolescents: Influences from variations in context and measure estimation. *PLOS ONE*, 9(7), e101356. <https://doi.org/10.1371/journal.pone.0101356>
- Koob, V., Mackenzie, I., Ulrich, R., Leuthold, H., & Janczyk, M. (2023). The role of task-relevant and task-irrelevant information in congruency sequence effects: Applying the diffusion model for conflict tasks. *Cognitive Psychology*, 140, 101528. <https://doi.org/10.1016/j.cogpsych.2022.101528>
- Kornblum, S. (1992). *Dimensional overlap and dimensional relevance in stimulus–response and stimulus–stimulus compatibility*. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior II* (pp. 743–777). Elsevier Science Publishers.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility—A model and taxonomy. *Psychological Review*, 97(2), 253–270. <https://doi.org/10.1037/0033-295X.97.2.253>
- Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, 38(5), 701–711. [https://doi.org/10.1016/S0028-3932\(99\)00106-2](https://doi.org/10.1016/S0028-3932(99)00106-2)
- Logan, G. D. (1985). On the ability to inhibit simple thoughts and actions: II. Stop-signal studies of repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 675–691. <https://doi.org/10.1037/0278-7393.11.1-4.675>

- Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E.-J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, 121(1), 66–95. <https://doi.org/10.1037/a0035230>
- Luna, B., Padmanabhan, A., & O’Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and Cognition*, 72(1), 101–113. <https://doi.org/10.1016/j.bandc.2009.08.005>
- Luo, C., & Proctor, R. W. (2022). A diffusion model for the congruency sequence effect. *Psychonomic Bulletin & Review*, 29(6), 2034–2051. <https://doi.org/10.3758/s13423-022-02119-8>
- Luo, J., Hao, C., Ma, N., & Wang, L. (2024). Sleep deprivation affects interference control: A diffusion model analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 50(2), 193–215. <https://doi.org/10.1037/xhp0001180>
- Martin, F., Long, S., Haywood, K., & Xie, K. (2025). Digital distractions in education: A systematic review of research on causes, consequences and prevention strategies. *Educational Technology Research and Development*, 73(6), 3423–3451. <https://doi.org/10.1007/s11423-025-10550-6>
- Masyn, K. E., Henderson, C. E., & Greenbaum, P. E. (2010). Exploring the latent structures of psychological constructs in social development using the dimensional-categorical spectrum. *Social Development*, 19(3), 470–493. <https://doi.org/10.1111/j.1467-9507.2009.00573.x>
- Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, 16(5), 798–817. <https://doi.org/10.3758/PBR.16.5.798>

- Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review*, 20(5), 819–858. <https://doi.org/10.3758/s13423-013-0404-5>
- 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>
- Mittelstädt, V., Miller, J., Leuthold, H., Mackenzie, I. G., & Ulrich, R. (2022). The time-course of distractor-based activation modulates effects of speed-accuracy tradeoffs in conflict tasks. *Psychonomic Bulletin & Review*, 29(3), 837–854. <https://doi.org/10.3758/s13423-021-02003-x>
- Noël, X., Van der Linden, M., Brevers, D., Campanella, S., Verbanck, P., Hanak, C., Kornreich, C., & Verbruggen, F. (2013). Separating intentional inhibition of prepotent responses and resistance to proactive interference in alcohol-dependent individuals. *Drug and Alcohol Dependence*, 128(3), 200–205. <https://doi.org/10.1016/j.drugalcdep.2012.08.021>
- Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, 106(37), 15583–15587. <https://doi.org/10.1073/pnas.0903620106>
- Patel, P. K., Green, M. F., Barch, D., & Wynn, J. K. (2025). Mechanisms and correlates of incentivized response inhibition in schizophrenia and bipolar disorder. *Journal of Psychiatric Research*, 183, 282–288. <https://doi.org/10.1016/j.jpsychires.2025.02.027>
- Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention*,

- Perception, & Psychophysics*, 72(7), 2013–2025.  
<https://doi.org/10.3758/APP.72.7.2013>
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108.  
<https://doi.org/10.1037/0033-295X.85.2.59>
- Ratcliff, R. (2013). Parameter variability and distributional assumptions in the diffusion model. *Psychological Review*, 120(1), 281–292. <https://doi.org/10.1037/a0030775>
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.  
<https://doi.org/10.1162/neco.2008.12-06-420>
- Ratcliff, R., Smith, P. L., & McKoon, G. (2015). Modeling regularities in response time and accuracy data with the diffusion model. *Current Directions in Psychological Science*, 24(6), 458–470. <https://doi.org/10.1177/0963721415596228>
- Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(4), 501–526.  
<https://doi.org/10.1037/xlm0000450>
- Schaefer, S. B., Radev, S. T., Göttmann, J., & Schubert, A.-L. (2026). Amortized Bayesian workflow for modeling congruency effects using the diffusion model for conflict tasks. *Computational Brain & Behavior*. <https://doi.org/10.1007/s42113-026-00266-y>
- Servant, M., & Evans, N. J. (2020). A diffusion model analysis of the effects of aging in the Flanker Task. *Psychology and Aging*, 35(6), 831–849.  
<https://doi.org/10.1037/pag0000546>
- Servant, M., Montagnini, A., & Burle, B. (2014). Conflict tasks and the diffusion framework: Insight in model constraints based on psychological laws. *Cognitive Psychology*, 72, 162–195. <https://doi.org/10.1016/j.cogpsych.2014.03.002>

- Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81(1), 174–176. <https://doi.org/10.1037/h0027448>
- Stanley, T. D., Doucouliagos, H., Maier, M., & Bartoš, F. (2024). Correcting bias in the meta-analysis of correlations. *Psychological Methods*. <https://doi.org/10.1037/met0000662>
- 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>
- Toh, W. X., & Yang, H. (2024). Do cognitive resources matter for eudaimonia? The role of inhibitory control in psychological well-being. *Trends in Psychology*, 32(2), 525–540. <https://doi.org/10.1007/s43076-022-00193-4>
- Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2015). Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology*, 78, 148–174. <https://doi.org/10.1016/j.cogpsych.2015.02.005>
- Valle-Inclán, F. (1996). The locus of interference in the Simon effect: An ERP study. *Biological Psychology*, 43(2), 147–162. [https://doi.org/10.1016/0301-0511\(95\)05181-3](https://doi.org/10.1016/0301-0511(95)05181-3)
- Van Den Wildenberg, W. P. M., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: A review. *Frontiers in Human Neuroscience*, 4. <https://doi.org/10.3389/fnhum.2010.00222>
- van Ravenzwaaij, D., Dutilh, G., & Wagenmakers, E.-J. (2012). A diffusion model decomposition of the effects of alcohol on perceptual decision making. *Psychopharmacology*, 219(4), 1017–1025. <https://doi.org/10.1007/s00213-011-2435-9>

- van Velzen, L. S., Vriend, C., de Wit, S. J., & van den Heuvel, O. A. (2014). Response inhibition and interference control in obsessive–compulsive spectrum disorders. *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00419>
- Verbruggen, F., McLaren, I. P. L., & Chambers, C. D. (2014). Banishing the control homunculi in studies of action control and behavior change. *Perspectives on Psychological Science*, 9(5), 497–524. <https://doi.org/10.1177/1745691614526414>
- Vermeulen, L., Braem, S., Ivanchei, I. I., Desender, K., García-Román, J. M., González-García, C., Ruz, M., & Notebaert, W. (2025). The temporal dynamics of metacognitive experiences track rational adaptations in task performance. *Communications Psychology*, 3(1), 96. <https://doi.org/10.1038/s44271-025-00282-x>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3). <https://doi.org/10.18637/jss.v036.i03>
- Weigard, A. S., Molloy, M. F., Sripada, C., & Heathcote, A. (2025). *The diffusion model's drift rate parameter primarily reflects efficiency, rather than speed, of evidence accumulation*. PsyArXiv. [https://doi.org/10.31234/osf.io/nu9gm\\_v1](https://doi.org/10.31234/osf.io/nu9gm_v1)
- Westfall, D. R., Gejl, A. K., Tarp, J., Wedderkopp, N., Kramer, A. F., Hillman, C. H., & Bugge, A. (2018). Associations between aerobic fitness and cognitive control in adolescents. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01298>
- White, C. N., Ratcliff, R., & Starns, J. S. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive Psychology*, 63(4), 210–238. <https://doi.org/10.1016/j.cogpsych.2011.08.001>
- White, C. N., Servant, M., & Logan, G. D. (2018). Testing the validity of conflict drift-diffusion models for use in estimating cognitive processes: A parameter-recovery study. *Psychonomic Bulletin & Review*, 25(1), 286–301. <https://doi.org/10.3758/s13423-017-1271-2>

Wylie, S. A., Ridderinkhof, K. R., Eckerle, M. K., & Manning, C. A. (2007). Inefficient response inhibition in individuals with mild cognitive impairment. *Neuropsychologia*, 45(7), 1408–1419. <https://doi.org/10.1016/j.neuropsychologia.2006.11.003>



