

KU LEUVEN

FACULTY OF PSYCHOLOGY AND  
EDUCATIONAL SCIENCES

**Metacognitive Insight into Interfering  
Information During Decision-Making Processes**

Master's thesis submitted for the  
degree of Master of Science in  
Master of Psychology: Theory and  
Research by:

**Eren Aslan**

Promotor: Prof. Dr Kobe Desender  
Co-supervisor: Dr Yvonne F. Visser  
Co-supervisor: Dr Luc Vermeulen

2025-2026



## Summary

The ability to control and suppress unwanted influences on decision-making processes is essential for adaptive behaviour. While research has shown that people can adjust their behaviour following conflict between two competing response alternatives, it remains unclear whether individuals have metacognitive insight into the extent to which conflicting and irrelevant information affects their decisions. Insight into interfering information can facilitate the optimisation of cognitive control strategies during the decision process.

This study aimed to understand whether subjective experiences of interference align with objective conflict modulation and whether these experiences correspond to behavioural performance and conflict levels induced by task manipulations. Fifty-one participants completed a Flanker task in which stimulus congruency and flanker distance were manipulated. After each trial, participants rated the degree of interference on a continuous scale, obtaining a trial-by-trial measure of their metacognitive sense of conflict. Linear mixed-effects models revealed that reaction times, accuracy, and interference ratings were sensitive to congruency and flanker distance. By fitting the Diffusion Model for Conflict (DMC) to empirical data, we showed that flanker distance primarily modulates latent parameters that measure conflict strength. We assessed the relationship between objective and subjective measures of interference by regressing peak amplitude, a proxy of interference in the DMC, and stimulus congruency on subjective reports. To further understand this relationship, simulation techniques were leveraged to model the relationship between peak amplitude and subjective interference.

The regression model provided evidence that individuals have insight into interfering information during decision-making. Model simulations further supported this claim by generating results that closely replicated the qualitative features of the empirical data. These findings support an account of conflict monitoring theory that bridges behavioural performance, computational modelling and subjective experience. Future research should extend these findings to other conflict tasks and incorporate neuroimaging techniques to further validate the results. Ultimately, combining subjective reports and objective measures of interference offers a more comprehensive approach to understanding adaptive decision-making when interfering information is present.

## Words of Appreciation

Throughout my academic career, I have quickly learned that no single victory can be considered mine only. If anything, I would not be standing here today if it were not for a never-ending list of amazing people who have aided, carried, tolerated, and supported me from the day I was born to this day.

Starting with two amazing parents, siblings and their spouses who were always there for me in times of need. My father (Hüseyin Aslan) and mother (Sevda Aslan) always knew how to fill me with confidence and tell me exactly what I needed to hear, even if I did not understand it at the moment. They were always my greatest supporters, and I will never cease to be grateful for every single opportunity they gave me. My siblings (Emel Aslan and Alperen Aslan) and their spouses (Mustafa Karademir and Büşra Anıl) respectively never stopped wondering whether their little brother was doing okay and always made sure to check up on me, even if I thought it was a nuisance. They collectively made me believe I can achieve anything I put my mind to, and I plan to do exactly that as long as they keep believing in me.

Next to my family, I had an amazing circle of close friends that made my day-to-day feel exciting and vibrant. Finding support in each other during times of struggle has been one of the most important aspects that has kept me going over the last two years. I want to thank them for their patience with me and my absence as a friend during these intensive studies. My heart goes out to my partner (Leylan Öykü Etkin), who for the last 1.5 years has been the closest to me and has always supported me through thick and thin. During this time, she has listened to thousands of thoughts regarding my thesis and always made sure I had someone to talk to.

Finally, I cannot express my gratitude enough to my supervisors during this period: Prof. Dr Kobe Desender, Dr Yvonne F. Visser and Dr Luc Vermeulen. Kobe has been incredibly patient and supportive and has challenged me to become the best researcher I could be over the last two years. His guidance has been pivotal in making this thesis possible, and he always made sure I could reach him in case of need. Yvonne is and will probably be one of the most helpful and kind people I will ever meet in academia. She always made sure I had someone to ask (an immense amount) of questions to and was always willing to help me. Her contribution to this thesis is invaluable, and I aim to be in her position one day, where I can similarly help master students as she helped me. Luc's support and contributions have been pivotal in understanding the more theoretical aspects of this thesis, and he was willing to aid me whenever I needed help.

Submitting this thesis was a victory, and I share it with all these amazing people. I thank all of you for being a part of this journey.

## Statement of Contribution

In the first week of starting this master's thesis, we held a meeting where we, as students, could choose a topic within Professor Desender's research project. I found the topic of metacognitive insight into interference interesting, and I spent the first weeks reading literature on the theoretical background of diffusion modelling, cognitive control, and metacognition.

After several meetings with Professor Desender, Dr Visser, and Dr Vermeulen, we decided on the experimental paradigm, which I programmed independently with regular feedback from my promotors. I conducted a pilot study with four participants and, after my supervisors' approval, began data collection independently.

In the meantime, I worked on the introduction of this thesis and got regular feedback from Professor Desender and Dr Visser. In total, my introduction was revised four times and was finished at the end of the first graduate year.

Subsequently, I began planning the data analysis and the project's steps and wrote the methodology. The behavioural analyses were based on Professor Desender's default workflow for linear mixed-effects modelling. I worked on this on my own for sustained periods and again got feedback from my supervisors where needed.

The drift-diffusion modelling part was extensive and took a long time to fine-tune. While I tried to take the lead where possible, my supervisors provided and fine-tuned the diffusion model — something I had no experience with — as needed. Particularly, the DMC was designed and implemented by Dr Vermeulen. During this time, I ran tests and parameter estimations with these models to gain some insight and fully understand the functions by the end of this project. I designed and implemented the restricted model described in this thesis myself, based on theoretical feedback from Professor Desender and Dr Visser. The data analyses were repeated several times in response to feedback and critical comments from everyone involved with this thesis.

Finally, I wrote out the results, discussion and conclusion during the latter half of the second graduate year and sent the final draft to my supervisors for feedback.

**Statement of Generative Artificial Intelligence Usage**

Generative Artificial Intelligence (AI) has been used in this thesis to control for grammatical errors, as well as debugging and programming queries as a co-pilot. All ideas, findings, theoretical perspectives and references were based on literature reviews and empirical data, alongside the author's own contributions.

## Table of Contents

|          |                                                               |           |
|----------|---------------------------------------------------------------|-----------|
| <b>1</b> | <b><i>Introduction</i></b>                                    | <b>1</b>  |
| 1.1      | The Ubiquity of Decision-Making                               | 1         |
| 1.2      | How We Make Decisions                                         | 1         |
| 1.3      | Cognitive Control & Conflict                                  | 4         |
| 1.4      | Conflict Monitoring Theory                                    | 5         |
| 1.5      | Metacognition and Subjective Awareness in Conflict Adaptation | 6         |
| 1.6      | The Mind's Tug-of-War: Conflict, Control and DDMs             | 7         |
| 1.7      | Understanding Conflict Through Subjective Experiences         | 11        |
| 1.8      | Present Study                                                 | 13        |
| <b>2</b> | <b><i>Methods</i></b>                                         | <b>14</b> |
| 2.1      | Participants                                                  | 14        |
| 2.2      | Apparatus and Stimuli                                         | 14        |
| 2.3      | Experimental Design                                           | 15        |
| 2.4      | Data Analysis                                                 | 16        |
| <b>3</b> | <b><i>Results</i></b>                                         | <b>24</b> |
| 3.1      | Linear Mixed-Effects Modelling Results                        | 24        |
| 3.2      | Parameter Estimation                                          | 27        |
| 3.2.1    | Global Parameter Estimation                                   | 27        |
| 3.2.2    | Participant-Level Parameter Estimation                        | 29        |
| 3.3      | Peak Amplitude and Subjective Interference                    | 30        |
| 3.4      | Simulation Results                                            | 31        |
| <b>4</b> | <b><i>Discussion</i></b>                                      | <b>35</b> |
| 4.1      | Limitations and Directions for Future Research                | 40        |
| <b>5</b> | <b><i>Conclusion</i></b>                                      | <b>42</b> |
| <b>6</b> | <b><i>References</i></b>                                      | <b>44</b> |

## List of Figures

|                                                                       |           |
|-----------------------------------------------------------------------|-----------|
| <b>Figure 1: The DDM Parameters</b> -----                             | <b>2</b>  |
| <b>Figure 2: The DMC Parameters</b> -----                             | <b>10</b> |
| <b>Figure 3: Stimuli Sequence During Trials</b> -----                 | <b>16</b> |
| <b>Figure 4: Model A: Reaction Times</b> -----                        | <b>25</b> |
| <b>Figure 5: Model B: Accuracy</b> -----                              | <b>26</b> |
| <b>Figure 6: Model C: Interference Ratings</b> -----                  | <b>27</b> |
| <b>Figure 7: Delta Plots: Full vs. Restricted Model</b> -----         | <b>28</b> |
| <b>Figure 8: Empirical vs. Simulated Data Comparison</b> -----        | <b>28</b> |
| <b>Figure 9: Paired T-tests between Conditions</b> -----              | <b>29</b> |
| <b>Figure 10: Model D: Amplitude and Congruency Interaction</b> ----- | <b>31</b> |

# 1 Introduction

## 1.1 The Ubiquity of Decision-Making

Decision-making is an integral part of everyday life. From simple choices like what to wear for a meeting to higher-stakes ones like a career switch, we are constantly making decisions.

Understanding how decisions are made is vital, as it provides insight into how the mind processes and selects information to execute actions. The study of how we make decisions across different situations has been a central focus of cognitive psychology. To investigate decision-making processes in a controlled setting, researchers often rely on simple experimental paradigms to extract informative data.

A common approach and operationalisation are two-choice speeded tasks, where participants must quickly and accurately choose between two options after a stimulus is presented. These paradigms simulate real-life decision-making under time pressure and have become indispensable in decision-making research, as these paradigms are relatively easy to design and yield clear, measurable outcomes. Researchers can derive valuable information, such as reaction time (RTs) and choice accuracy, to evaluate the cognitive underpinnings of decisions. These variables might reveal, for example, how people emphasise accuracy over speed and vice versa. This information subsequently helps us understand how people make errors in high-stakes situations, such as in air traffic control or emergency care.

## 1.2 How We Make Decisions

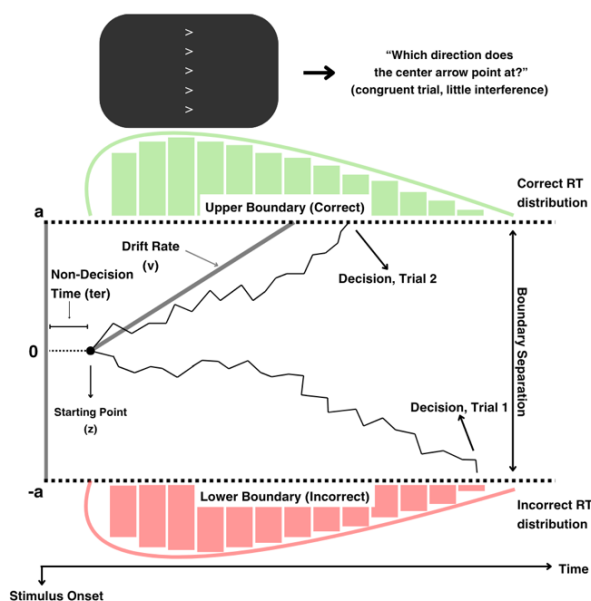
As decision-making constitutes an integral part of our lives, the question of *how* decisions are made and guided becomes prominent. Sequential sampling models (SSMs) are a class of models that have been successful in formalising simple decision-making processes. Sequential sampling is based on the idea that decisions are made through an evidence-accumulation process in which noisy evidence is collected until a threshold is reached, at which point a decision is made (Ratcliff et al., 2016). When a decision is formed by weighing the evidence between options — A has more evidence than B — the evidence accumulation process is described as a random walk model: Noisy evidence accumulates in a semi-random manner (Ratcliff et al., 2016). SSMs propose that, instead of assuming decisions are made instantly, the cognitive system takes competing choices into account dynamically

and in a time-dependent manner. This framework can account for differences in accuracy and RT distributions, offering insights previously not obtainable.

One influential class of SSMs, the Drift Diffusion Model (DDM), has become canonical for understanding decision-making processes. This class of models conceptualises two-alternative choice tasks as a continuous process of noisy evidence accumulation, modelling both the accumulation and the time to reach a decision (Ratcliff et al., 2016). According to Ratcliff and colleagues (2016), the DDM dissects the decision-making process into distinct latent parameters (as shown in Figure 1): the *drift rate* represents the average rate at which information is gathered in favour of either choice. It can be considered the average rate and quality of the evidence accumulation process and is assumed to be constant in the conventional DDM. *Boundary separation/threshold* reflects the amount of evidence required to reach either decisional threshold. A higher threshold (or wider boundary separation) indicates the need for more evidence: the decision will, in this case, be slower but more accurate. *Non-decision time* describes processes that surround (and are unrelated to) the decision itself, such as perceptual encoding and motor responses. *Starting point* (also referred to as bias) represents the initial position of the evidence accumulation process relative to the boundaries, and indicates whether there is a bias towards either choice.

### Figure 1

*The DDM Parameters. The first trial leads to an incorrect response, while the second trial leads to a correct response*



The DDM has been influential in decision-making literature and provides several key advantages. Speed/accuracy trade-offs can be disentangled, offering more accurate estimates of evidence quality (rather than just accuracy and RTs alone). Individual differences can be picked up more easily because the model handles variability in RTs and accuracy better than simpler approaches. In general, the model is flexible and broadly applicable, used in a multitude of task settings (Ratcliff et al., 2016).

Applications of the DDM outside cognitive and mathematical fields can be found in, among others, clinical psychology. Huang and colleagues (2015) investigated how dopaminergic medication influences decision-making in Parkinson's disease (PD) patients. Using a dot-motion task with varying difficulty and speed-accuracy requirements, drift diffusion modelling revealed that PD patients on dopaminergic medication had lower drift rates and made more errors in difficult tasks, suggesting an impairment in perceptual evidence accumulation. In contrast, differences in non-decision time and performance, alongside evidence that dopaminergic medication did not induce impulsive decisions under speed pressure, suggest that medication primarily affected motor-related processes rather than the decision-making process itself. Using a go/no-go task, Huang-Pollock et al. (2016) challenged traditional ideas of evidence accumulation in children with ADHD and showed that fast-paced tasks may in fact impair the evidence accumulation process, not enhance it. Moustafa et al. (2015) used a classification task and drift diffusion modelling to examine schizophrenic patients' decision-making processes compared to healthy controls. Patients with schizophrenia showed slower non-decision time, more caution when giving a response and deficits in learning from punishment. These findings suggest that accuracy is prioritized over speed and highlight the DDM as a tool to better understand cognitive deficits.

Interesting to note is that the DDM — and sequential sampling methods more broadly — are not merely conceptual and mathematical approximations of decisions. Rather, researchers managed to relate these models to concrete neural activity. Kelly and O'Connell (2013) probed how both internal and external factors affect the dynamics of the decision variable using a modified dot-motion task. Using electroencephalography (EEG), they measured neural signals while participants made decisions during a random dot-motion task — a speeded two-choice task where a cloud of dots is presented with

a certain percentage moving coherently in either direction — and found that the centroparietal positive potential (CPP) reflected a supramodal decision signal. The CPP reproduced two typical sequential sampling processes: the CPP buildup scaled consistently with evidence strength and reached a consistent threshold when a decision was performed, and the variations in CPP buildup predicted RTs, suggesting that neural states can modulate the speed and accuracy of decisions. This study provides a basis for the DDM to be understood not only as a mathematically elegant abstraction, but also a framework that can be associated with neural correlates.

### 1.3 Cognitive Control & Conflict

Even though decisions are made constantly — and models such as the DDM allow us to better understand them — not all decisions are straightforward. Both in real life and experimental tasks, we face conflicting and distracting information that interferes with the decision process. Consider the following scenario: while waiting at a stop light in the car, the light turns green. However, there is a traffic officer signalling that it is forbidden to continue. For a second, one might feel the urge to press the gas pedal (due to the green light) but manages to control this situation (due to the officer's signal). Situations like this require cognitive mechanisms to suppress the primary response in exchange for a contextually goal-relevant one.

This ability to change behaviour and thought processes by selectively managing attention and responses based on tasks is commonly known as cognitive control (Botvinick et al., 2001). Cognitive control allows the perceptual system to align decisions with active goals rather than irrelevant and distracting ones. In an experimental setting, these conflicting situations can be operationalised in several ways.

The Stroop task is a classic example, where one must name the colour of a word while ignoring its semantic meaning, such as the word “blue” written in red (Stroop, 1935). Another example is the Eriksen-flanker task, where a target stimulus such as an arrow is displayed along with distracting flankers that can be either congruent or incongruent with the target (Eriksen & Eriksen, 1974). Results of these tasks consistently show faster RTs and smaller errors in congruent rather than incongruent trials, often referred to as **congruency effects** (Stroop, 1935; Eriksen & Eriksen, 1974).

Plenty of research has replicated these results, along with observing other effects, particularly on how people adapt their behaviour after facing conflict. Gratton and colleagues (1992) showed during a letter flanker task that the magnitude of congruency effects tends to be smaller following incongruent trials, a pattern called the **Gratton effect**. This effect is thought to occur because cognitive control mechanisms bolster after experiencing conflict. Control systems become focused after incompatible trials, reducing the effects of distractors in subsequent trials (Botvinick et al., 2001). The Gratton effect thereby reflects a reactive trial-by-trial adjustment of strategy as a response to conflict. A frequently used term to refer to these general patterns, where performance is dependent on previous trial congruency, is Congruency Sequence Effects (CSEs; Egner, 2007). The presence of these conflict-related effects shows not only the influence of task-irrelevant cues on the decision process, but also that the cognitive system keeps track of performance and allocates control when needed on a trial-by-trial basis.

#### **1.4 Conflict Monitoring Theory**

A foundational framework in the domain of conflict-related adaptation is the conflict monitoring theory. This theory accounts for how the cognitive system discerns conflict between competing stimuli and uses this information to initiate adjustments in cognitive control, allowing performance to be more efficient in these situations (Botvinick et al., 2001). This compensatory system deals with the inadequacy of current processes and the bias towards goal-irrelevant cues. According to this theory, the cognitive system constantly monitors for the presence of conflict. When conflict is identified, for example during an incongruent trial, an increase in control is triggered. This increase in control causes the system to focus more on task-relevant stimuli rather than distractors.

Conflict monitoring theory has been crucial in explaining phenomena such as CSEs: When the cognitive system — that continuously monitors for conflict — detects an incongruent trial, conflict monitoring is increased in subsequent trials. This allows for enhanced focus on the target, while filtering out distractors. This process results in less interference in the next trial (Botvinick et al., 2001; Gratton et al., 1992). The system's monitoring of conflict employs changes that allow for more efficient performance after experiencing conflict, creating observed CSEs (Egner, 2007). This

highlights the importance of continuous performance monitoring in decision-making tasks involving conflict.

### **1.5 Metacognition and Subjective Awareness in Conflict Adaptation**

Cognitive control has long been studied based on the premise of objective and computational measures (Ratcliff et al., 2016). While these perspectives have gained a lot of attention and yielded powerful insights, they provide a limited account of our capacity to reflect on and be aware of decision-making processes as they occur. Metacognitive abilities — our ability to purposefully be aware of, select, evaluate and control cognitive functions — provide a complementary perspective alongside objective measures. In this context, metacognition emerges as awareness of one's performance while making a choice, allowing people to sense when they experience conflict, interpret subjective experiences, and relate these to behavioural regulation. Studying metacognition in cognitive control therefore offers an integrative account that connects subjective experience with control dynamics, providing further insight into the role of cognitive control in decision-making processes.

To illustrate, Desender and colleagues (2014) showed that subjective awareness of conflict alone can initiate adaptive cognitive control processes. In a subliminal priming task (where stimuli are presented unconsciously to influence subsequent trials), a prime arrow pointing either left or right was briefly presented and masked, rendering it unnoticeable. The prime was followed by a target arrow, and participants were instructed to choose the direction of the arrow quickly and accurately. After each trial, participants indicated whether they felt a conflict between prime and target, choosing between four graded options (ranging from *"I think there was a conflict"* to *"I think there was no conflict"*). Although participants could not consciously perceive the primes, they still reported a sense of conflict more often on incongruent trials than on congruent trials. The Gratton effect only emerged when participants had experienced conflict on the previous trial. These findings indicate that individuals can experience an introspective sense of conflict even when the source of that conflict is not consciously perceived, and that this subjective experience is functionally captured in metacognitive judgements (Desender et al., 2014).

Maniscalco and colleagues (2012) similarly found, using a modified flanker task, that the subjective perception of distracting flankers heightened interference effects during the task, further

supporting the notion that cognitive control is not solely stimulus-driven but is also shaped by subjective judgements.

Although subjective experiences are distinct from objective stimulus properties such as congruency, both appear to contribute to the regulation of cognitive control. Questienne and colleagues (2021) argue that this subjective awareness plays a functional role in cognitive control mechanisms, with several factors — including reaction times, motor hesitation, and visual conflict — jointly shaping how interference is experienced. Extending this further, Bognar and colleagues (2024) showed that cognitive control mechanisms engage gradually and continuously in proportion to the amount of conflict subjectively experienced, suggesting that control is not a binary on-off switch but scales with the felt intensity of interference.

Together, this body of work converges on the idea that subjective experiences of interference are constructed from multiple observable cues that are indicative of task performance such as response speed, congruency, and response dynamics (Desender et al., 2017; Questienne et al., 2021). However, this raises a deeper question: are these reports merely post-hoc inferences built from surface-level cues, or do they reflect something more direct — a genuine read-out of the internal processes that are actually driving the decision? As the aforementioned studies established that observable cues shape subjective reports, the question of whether these reports reflect the underlying decision process itself remains unanswered.

## **1.6 The Mind's Tug-of-War: Conflict, Control and DDMs**

Considering the aforementioned advantages of using the DDM as a model for decision-making, the intriguing question of whether this framework can also be used to better understand computational mechanisms underlying decision performance in conflict tasks arises. As previously noted, the standard DDM assumes a constant drift rate and singular process of evidence accumulation. A conflict task (such as the Flanker or Stroop task) is more complex: it introduces two sources of evidence: a target-related source and a distractor-related source. The strategies implemented in a standard DDM paradigm might not be optimal when conflict is introduced in the task (Servant et al., 2014). The decision process is thus further complicated by competing evidence accumulation processes, dynamic changes in drift rate as participants shift their focus from distractors to targets and

strategic adjustments to boundaries while making a decision (Ulrich et al., 2015) — aspects due to which standard DDMs fail to capture the decision process as a whole during conflict paradigms.

An example revealing the shortcomings of the standard DDM relating to conflict is its inability to fully capture **delta functions**, which provide a simple way to understand how congruency effects change for different reaction times. Delta plots display, for each RT quantile bin, the difference in RT between incongruent and congruent trials (y-axis) against the mean RT across both trial types at that quantile (x-axis), thereby showing how the congruency effect evolves across the response time distribution. Delta functions with a positive slope are often related to increasing congruency effects, while negative-going delta functions reflect diminishing congruency effects — the latter indicating that individuals are able to suppress interfering information given enough time (Ulrich et al., 2015). Delta plots have become crucial to study conflict in tasks, and the DDM struggles to account for such complex distributional patterns and is not able to generate negative-going delta plots (Servant et al., 2014; Ulrich et al., 2015).

The fact remains that the standard DDM cannot fully account for complexities that arise regarding conflict. In order to capture conflict in decision-making tasks, the DDM framework needs extension to account for non-linear evidence accumulation and faithfully understand the decision processes. Below, some of these alternative frameworks and extensions will be discussed.

An example by Hübner et al. (2010) is the **dual-stage two-phase model** (DSTP). The model addresses the need to control response mechanisms to stimuli by incorporating two stages of stimulus selection: an early stage which serves as a perceptual filter with low selectivity and a later, more selective stage which functions on a conceptual level. Hübner and colleagues (2010) showed that the fit of the DSTP model on their data surpassed the fit of more traditional single-stage models.

Another well-established framework is the **Shrinking Spotlight Model**, an expansion of traditional spotlight theories, where attention is thought of as a focal point in the visual field that enhances perception. In the Shrinking Spotlight Model, the attentional frame is initially broad and allows for interference. The locus of attention slowly narrows as attention becomes more engaged, reducing the influence of distractors in the process (White et al., 2011). This model, analogous to a zoom lens, introduced the idea that attention can dynamically change its focus as needed.

To better understand diffusion models in conflict situations, Servant et al. (2014) examined two known phenomena: **Piéron's law** and the **Wagenmakers-Brown** law across different conflict tasks. The first describes how RTs decrease as stimulus intensity increases, and the second characterises a positive linear relationship between mean RT and RT variance. Their results showed that while Piéron's law held in both congruent and incongruent trials, congruency systematically affected the Wagenmakers-Brown's law: incongruent trials were consistently associated with a reduction in RT variability, with smaller observed standard deviations in comparison to congruent trials (Servant et al., 2014). To interpret these findings, Servant and colleagues (2014) turned to modelling. Here, two extended DDM models (reflecting DSTP and Shrinking Spotlight Model) were used, allowing the models to differentiate between targets and distractors. These findings showed that mechanisms underlying conflict effects are independent of the mechanisms regarding target processing (Servant et al., 2014).

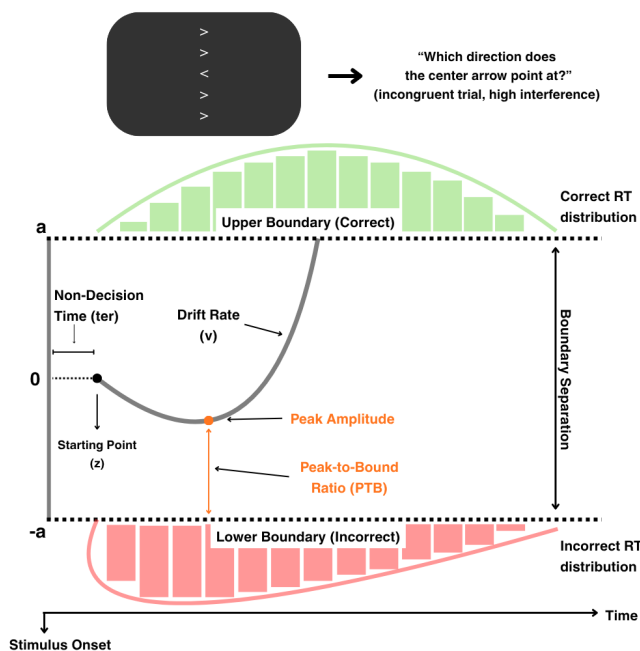
Another pioneering example in computationally modelling conflict in decision tasks is the **Diffusion Model for Conflict Tasks (DMC)**, proposed by Ulrich et al. (2015). At its core, the DMC proposes that two distinct processes simultaneously drive the decision process: a *controlled* process that tracks task-relevant information, and an *automatic* process triggered by task-irrelevant information. Both processes feed into a superimposed evidence accumulator that determines which decision is made. The controlled process contributes a constant drift of evidence toward the correct response throughout the entire trial — like a traffic officer patiently signalling you to wait, exerting a reliable influence. The automatic process exerts a transient influence on the decision process, like the reflexive urge to press the gas pedal the moment a traffic light turns green. Mathematically, this transient impulse is captured by a **Gamma distribution**, whose shape and scale parameters determine how sharply it rises and how quickly it subsides — effectively setting the time scale over which conflict plays out.

In an experimental paradigm, the direction of the automatic process depends on the congruency of stimuli on a trial. On incongruent trials, where the distractor points away from the correct response, the automatic process pushes the accumulator in the wrong direction — temporarily slowing or even reversing the build-up of evidence toward a correct response. On congruent trials, the

automatic process aligns with the correct response, adding an early boost of evidence on top of the controlled drift and speeding up responding. The peak amplitude of the automatic process — the highest point of the Gamma curve — serves as the key index of how strongly the distractor captures the decision, also referred to as the degree of irrelevant capture (Vermeulen et al., 2025). A large peak amplitude means stronger interference on incongruent trials and stronger facilitation on congruent ones. The DMC, through its architecture, can account for different phenomena simpler models cannot. It can naturally model negative-going delta plots, and accounts for distributional effects and accuracy patterns of conflict tasks such as the flanker task (Ulrich et al., 2015).

## Figure 2

*The DMC parameters. In this example trial, the mean drift rate is pulled toward the incorrect boundary. The DMC can account for fast errors due to interference better compared to the DDM.*



This wave of extended diffusion models has profoundly advanced the theoretical understanding of cognitive control systems and respective adjustments. From now, we will refer to these models as conflict drift diffusion models (**CDDMs**): models that account for patterns observed in conflict tasks where standard DDM models fall short.

Fischer and colleagues (2018) were able to put this knowledge and the advancements of CDDMs in practice: Their goal was to better understand exactly how people adapt decisions post-error and link these findings to neural data using EEG. Participants performed an Eriksen-flanker task

where flanker-target distance was manipulated. Using a multi-stage DDM (an implementation of CDDMs that separately weights distractor and target evidence), they compared trial-wise evolutions of decision parameters with lateralised beta-band power. Their research identified several independent changes model parameters post-error: (1) boundary separation was wider, indicating more caution, (2) signal-to-noise ratio decreased due to the suppression of distracting stimuli and (3) drift rate was momentarily slowed down, most likely indicating a disturbance in focus. These elements combined overall benefited subsequent trials: the likelihood of repeating an error was reduced. Crucially, Fischer et al. (2018) also related these models to EEG beta-band activity, showing neural correlates in sensorimotor cortices of these decision-making adjustments — specifically, that lateralised beta power tracked the evolving decision variable and reflected both the increased response threshold and the reduced influence of distractors following errors. Similar to the results of Kelly and O'Connell (2013), these findings show that mechanisms engaging in conflict modulation track closely with real-time neural signals rather than existing solely as mathematical abstractions. The convergence between model parameters and beta-band lateralisation suggests that conflict-sensitive diffusion models capture processes with a genuine neurophysiological basis.

## **1.7 Understanding Conflict Through Subjective Experiences**

Having established that the presence of interfering information necessitates theoretical extensions beyond the standard DDM through the advancements of CDDMs — and that individuals demonstrate that conflict is experienced and use this awareness to regulate performance, it becomes essential to understand whether subjective awareness reflects underlying computational mechanisms described by CDDMs. Conflict, specifically within the DMC, is instantiated as a transient and automatic activation that influences evidence accumulation. However, it remains an open question to what extent subjective feelings of interference correspond to the latent dynamics of these model-based approaches.

Subjective experiences, such as metacognitive awareness, are not just cognitive byproducts but play a crucial role in conflict monitoring and regulation. These experiences can occur even without conscious awareness and drive behavioural adjustments. Important to note is that this awareness is separate from objective task features, yet it still contributes to regulating cognitive control. As follows,

the need for research on this topic becomes apparent: Cognitive control is not only shaped by external, objective task-related factors but also depends on how conflict is felt, perceived and interpreted by the individual. Insight into these subjective experiences may reveal more on the mechanisms of internal conflict monitoring and control.

The role of subjective experiences in performance adjustments was recently studied by Vermeylen et al. (2025). In their study, participants completed a flanker task while regularly reporting their feelings (fatigue, boredom, effort and others). Combining these subjective experiences with neural data and diffusion modelling, temporal dynamics of task performance and experience were measured. The results indicated that differences in diffusion-model parameters were closely related to the measured experiences (Vermeylen et al., 2025). These metacognitive experiences appeared to distinctly track different cognitive adaptations, indicating that fatigue, boredom, effort and other experiences might have a one-on-one link with particular control operations being activated. Simply, how a person feels about the task at hand plays a significant role in how the task is approached. If a person feels conflict and interference, one might increase control to reduce these effects, similarly to how people exhibit conflict avoidance in real life.

The emerging idea is that cognitive control mechanisms are not singular, and that the question is not whether control is active, but rather which specific aspect of control is changing (Fischer et al., 2018). If individuals can introspect on the degree to which irrelevant information is capturing their decision process, a person sensing strong automatic capture might selectively increase suppression of task-irrelevant information.

Consistent with this possibility, Desender and colleagues (2018) showed that subjective confidence predicts information-seeking behaviour, with participants more likely to request additional evidence before coming to a decision when confidence was low. This demonstrates again that metacognitive reports are not merely byproducts of the decision process but can actively shape subsequent behaviour — suggesting that awareness of one's internal state, where present, may support more flexible and goal-directed control. These theoretical developments raise the question of whether people have metacognitive insight into their subjective experiences during conflict situations, and if so, what internal processes those experiences actually reflect.

## 1.8 Present Study

Based on the previous literature and findings, one fact becomes clear: Decision-making, cognitive control and metacognitive insight are deeply intertwined. These perspectives however have largely been studied in isolation: Cognitive control research documents how people behaviourally adapt to conflict, while metacognition research examines how people reflect on and report their experiences. What remains unknown is whether subjective interference reports reflect a direct read-out of the latent computational processes driving the decision.

This distinction matters for two reasons. First, it speaks to the nature of metacognition itself. Desender and colleagues (2017) demonstrated that subjective difficulty is constructed from multiple observable cues suggesting that individuals infer their subjective experiences from task performance. However, if people instead have direct access to the state of a specific latent parameter such as peak amplitude, their reports would be informationally richer, reflecting the strength of automatic capture rather than a surface-level proxy for it. Second, direct metacognitive access to parameter states would be far more adaptive from a control perspective. A person who can sense how strongly task-irrelevant information is pulling their decision could deploy targeted suppression in proportion to that pull — a more efficient strategy than responding uniformly to the mere presence of conflict.

We hypothesize that latent conflict parameters derived from the DMC — specifically, the peak amplitude of automatic activation — will predict self-reported ratings of interference. Greater automatic capture should correspond to stronger subjective experiences of interference on incongruent trials. Participants will perform an adapted Eriksen-flanker task, during which subjective ratings of interference will be collected. The DMC will be fitted to this data and resulting latent parameters will be further analysed. If latent conflict parameters such as peak amplitude are associated with subjective reports of interference, this would indicate that individuals have insight into conflict-related processes captured by the DMC. Additionally, we will use simulation techniques to generate behavioural data under the assumption that subjective interference ratings are a direct function of latent conflict parameters, particularly the amplitude of automatic activation. By comparing these simulated results with the empirical data, we can assess whether the observed pattern is consistent with a generative process in which subjective awareness tracks model-derived conflict strength.

## 2 Methods

### 2.1 Participants

Undergraduate psychology students from KU Leuven were recruited using the university's Experiment Management System. Students were rewarded with credits for an undergraduate course if they completed participation. Additionally, some participants were recruited by convenience sampling to meet the total quota. These additional participants were volunteers and did not gain anything by participating. All 51 participants had normal or corrected-to-normal vision, and the average age was 20.71 years ( $SD_{age} = 3.02$ ,  $min_{age} = 17$ ,  $max_{age} = 28$ ). Forty-four participants were right-handed, while the remaining seven were left-handed. Participants were informed that the study would involve partaking in a simple computer task without any health risks during a session that would last around 60 minutes. Information sheets and informed consent forms were distributed and signed prior to participation. Participants could withdraw at any time throughout the experiment. Throughout the whole experimental process, anonymity and confidentiality of participant information and data were assured.

An experimenter was present throughout the sessions, and participants were allowed to ask for help at any time. We verified participants' primary language by asking them directly, confirming it was either Dutch or English. Dutch-speaking participants indicated they had sufficient proficiency in English and were informed that they could request translations during the task, which was conducted in English. Informed consent forms and information sheets were available in both languages. This study was not preregistered due to the primarily explorative nature of the research question.

### 2.2 Apparatus and Stimuli

The task was conducted on a 1920 x 1080-pixel resolution LED monitor with a refresh rate of 60 Hz in a quiet, normally lit lab environment. Participants were instructed to position themselves about 60 cm from the monitor. A modified version of the Eriksen-flanker task was used and programmed using Python and the PsychoPy package (Eriksen & Eriksen, 1974; Peirce et al., 2019). All trials used white stimuli on a black background, presented in the default PsychoPy font. Participants could provide a response to the stimulus by using a standard computer mouse, clicking either left or right. The mouse cursor was not visible during trials. They then reported their subjective

feeling of interference after each trial by moving the mouse cursor over a continuous scale and clicking the space bar to confirm.

### **2.3 Experimental Design**

Our study aimed to partially replicate the results of Fischer et al. (2018) under similar conditions, and the experimental method was designed as such. A within-subject 2 x 2 factorial design was used with factors congruency (congruent or incongruent) and flanker distance (close or distant). The stimuli consisted of white arrow symbols (">" or "<") serving as either a target in the centre or flanker in the periphery, expressed in visual degrees. The arrows were 1.3° in height and 1.9° in width. Every trial, five arrows were shown in a vertical column: a central target and four flankers around it (two above and two below), spaced 1.75° apart in the close condition and 4° apart in the distant condition. In congruent trials, the target and flankers pointed in the same direction, while in incongruent trials the flankers pointed in the opposite direction of the target.

Prior to the main experiment, participants underwent a practice stage to familiarise themselves with the task. On-screen instructions emphasised the need to balance speed and accuracy, and participants were instructed to keep their eyes fixed at the centre of the screen throughout. The practice stage consisted of two introductory phases followed by a full practice phase. In the first introductory phase, a single central arrow was presented without flankers for 300 ms, allowing participants to acclimatise to the stimuli and response timing. In the second, the target was shown for 33 ms without distractors, giving participants practice at the full speed of target presentation. Both introductory phases consisted of two blocks of five trials each. The subsequent practice phase mimicked the main task — flanker arrows appeared for 83 ms before the 33 ms target presentation, after which all stimuli were removed — and consisted of two blocks of twenty trials each.

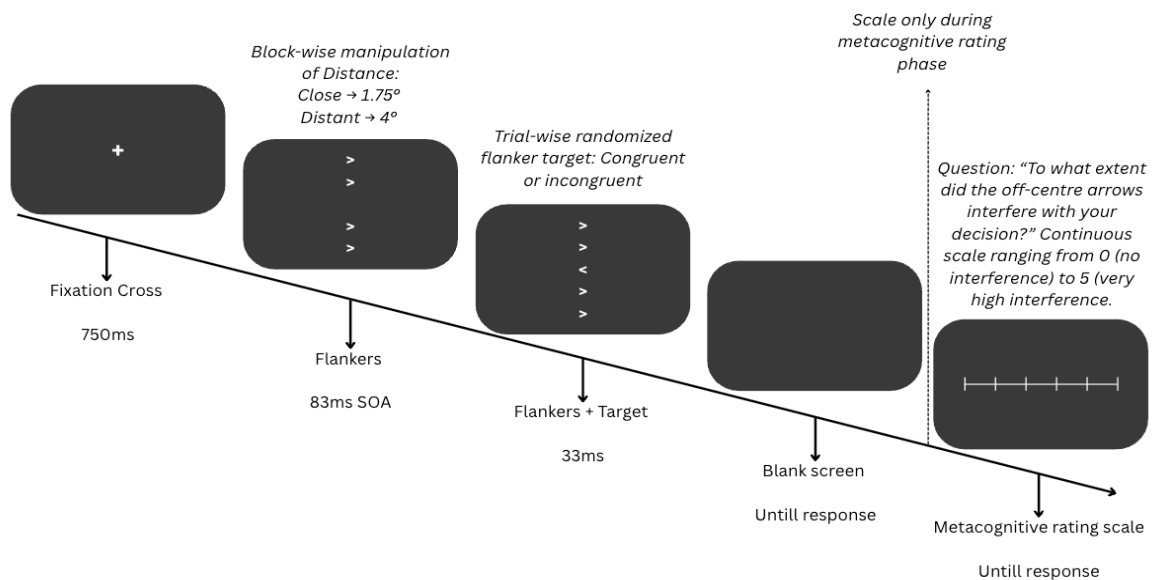
Each trial in the main practice phase began with a fixation cross displayed for 750 ms. During a brief stimulus-onset asynchrony (SOA), the four flanker arrows were presented for 83 ms, followed by the target arrow for 33 ms, after which all stimuli were removed. Participants indicated the direction of the target by clicking the respective mouse button, and the screen remained blank until a response was registered. Participants were encouraged to be both fast and accurate. Flanker distance was held constant within a block and alternated between blocks, ensuring equal exposure to both

distance conditions. Within each block, congruent and incongruent trials were randomised with equal probability. The main practice phase consisted of four blocks of fifty trials each. Participants received block-wise feedback on their accuracy and RT, which also served as a short break. If RT exceeded 400 ms, participants were instructed to respond faster in subsequent blocks: if accuracy fell below 100%, they were instructed to respond more accurately. Both messages could appear simultaneously if neither criterion was met.

In addition to the main practice phase, participants completed a metacognitive phase consisting of eight blocks of fifty trials each, in which a rating scale was incorporated immediately after each response. Participants rated how much the flanker arrows had interfered with their decision on that trial on a continuous scale from zero ("No Interference") to five ("Very High Interference"), using an on-screen slider confirmed with the space bar. Prior to this phase, participants were instructed that their subjective sense of interference does not necessarily correspond to objective stimulus properties, such as congruency — congruent trials might also come with a sense of interference — ensuring they relied on their own thoughts for rating the interference of each trial.

**Figure 3**

*Overview of Stimuli Sequence During Trials*



## 2.4 Data Analysis

All training phases, except for the metacognitive phase, were excluded from further data-analyses. Trials were excluded based on a lower bound of 200 ms to account for random responses and

an upper bound of 3000 ms based on visual inspection of the RT distribution. One participant was left out of subsequent analyses due to accuracies below chance level — this threshold was based on accuracy levels below 50% for a task with two alternatives. All data-analyses were conducted in RStudio using R (version 4.4.1).

**Behavioural Analyses.** Behavioural performance was analysed descriptively across the two within-subject factors flanker congruency and distance. Summary statistics were computed for three outcome measures: (1) reaction time, (2) accuracy and (3) subjective interference ratings. Accuracy was operationalised as the proportion of correct trials per condition. Given the natural right-skewed distribution of the RTs, median values were computed instead of means. These values were computed for each of the four condition combinations (congruent-close, incongruent-close, congruent-distant, incongruent-distant) on participant-level data and then grouped by condition to ensure equal weights of all participants on the final results. Interference ratings were calculated by taking the average value of all trials in each condition.

To formally test the effects of the within-subject factors on task performance and subjective ratings, we ran a series of linear mixed-effects model (LMM) analyses on the data. Due to the experimental setup, responses are nested within participants, and group sizes are unbalanced. This results in repeated measurements that are interdependent. Each participant contributed multiple data points for their responses, which are likely to correlate. Traditional linear regression cannot capture these model complexities. LMMs account for this interdependence by adding random effects that model variability for each participant and each predictor and can better handle unbalanced data. This method allows for the estimation of population-level fixed effects while taking individual differences into account. An overview with all models and their respective specifications can be found in Table 1.

All models included the same predictors: Stimulus distance, congruency and their interaction. The predictors were dummy coded beforehand to allow for more simple interpretation. Model A predicted RTs, testing the interference effects on speed. RTs were log-transformed and grand-mean centred before modelling to correct for the left-tailed skew and to improve model fit. Visual inspection of the diagnostic plots revealed mild violations of normality and homoscedasticity, consistent with typical RT distributions. Given the large sample and robustness of LMMs, no further transformations

were applied. Model B predicted accuracy: as the outcome variable accuracy is categorical rather than continuous in the latter model, a logistic mixed-effects model was utilized. Visual inspection of diagnostic plots indicated no major violations of model assumptions. Model C predicted the continuous subjective ratings of interference. Visual inspection of diagnostic plots indicated no major violations of model assumptions aside from minor deviations from normality, though these were not severe. No modifications were considered for this model.

Trials where participants answered incorrectly were left out of the analyses of models A and C. All trials, both correct and incorrect, were used in model B. These models were fitted using the lme4 (1.1.37) package. Random intercepts and slopes were modelled for all three predictors by each participant using a maximal random-effects structure. The significance of the fixed effects of models A and C was assessed with a Type III ANOVA test, obtaining  $F$  and  $p$ -values using the lmerTest (3.1.3) package.  $\chi^2$  tests and  $p$ -values were obtained for model B using the car (3.1.3) package.

Model diagnostics were assessed by visually checking error residuals for normality and diagnostic plots for homoscedasticity using the performance (0.15.2), ggResidpanel (0.3.0) and car (3.1.3) packages for the first two models and the DHARMa (0.4.7) package for the logistic model. Multicollinearity in the models was evaluated using variance inflation factors (VIF) for the fixed effects of models A and C and adjusted-VIF values to account for the categorical outcome variable in model B. For models revealing statistical significance of the main effects or interactions, pairwise contrasts were computed using the Tukey method to control for family-wise error rates. For ease of interpretation, model estimates were back-transformed to the original measurement scales (reaction time in seconds, accuracy as percentages, and interference scores on the 0-5 scale) and are reported with their 95% confidence intervals as Estimated Marginal Means (EMMs). EMMs are preferable to raw empirical averages because they account for random-effects structures of the models and any design imbalance. EMMs are also more directly interpretable than raw regression coefficients.

**DMC Parameter Estimation.** Participants' behavioural data were fitted using a custom version of the DMC implemented in C++ and analysed in R. A custom implementation was chosen over an established R package such as DMCFun (Ulrich et al., 2015) to allow for subsequent extensions of the model that go beyond the scope of this thesis. As specified in the DMC, decision formation was

described as a sequential sampling process combining a controlled drift rate and a time-varying automatic activation component. The evolution of accumulated evidence  $x(t)$  follows a stochastic differential equation:

$$dx(t) = [v_c + c \cdot v_a(t)]dt + s dW, \quad (1)$$

where  $v_c$  is the constant controlled drift rate,  $c \in \{-1, 1\}$  codes the congruency condition (incongruent vs. congruent),  $s = 4$  is the diffusion coefficient (fixed), and  $v_a(t)$  is the automatic activation pulse.

As previously discussed, the automatic component was implemented as a gamma-shaped pulse governed by a shape parameter  $\alpha = 1$  and a time scale parameter  $\tau$  (in milliseconds). The gamma pulse is defined as the derivative of a scaled gamma density, following Ulrich et al. (2015):

$$v_a(t) = A \cdot \frac{e^{-t/\tau}}{\tau} \cdot \left( \frac{t e}{(\alpha - 1)\tau} \right)^{\alpha-1} \left( \frac{\alpha - 1}{t} - \frac{1}{\tau} \right), \quad (2)$$

where  $A$  is the peak amplitude. Congruency determines the direction of the automatic activation via the sign factor  $c$  multiplying  $v_a(t)$ . The decision process starts at  $x(0) = 0$  and terminates when the accumulated evidence reaches an upper boundary  $a$  (correct response) or a lower boundary  $-a$  (incorrect response). RT is the sum of the decision time and a non-decision time component  $t_{er}$ . The model further included two sources of inter-trial variability. Starting point variability was implemented by drawing the initial evidence from a uniform distribution scaled by a Beta distribution:

$$x(0) \sim 2a \cdot (Beta(\beta, \beta) - 0.5), \quad (3)$$

with shape parameter  $\beta > 0$  (when  $\beta = 0$  the starting point is fixed at 0). Non-decision time variability was implemented by drawing  $t_{er}$  independently for each trial from a normal distribution with mean  $\mu_{ter}$  and standard deviation  $\sigma_{ter}$ .

The diffusion process was approximated numerically with a fixed time step of 1 ms. Gaussian random numbers for the diffusion noise were generated via the RcppZiggurat package (0.1.8). For each evaluation of the model, 5 000 trials were simulated (2 500 congruent, 2 500 incongruent).

Parameters were estimated using a differential evolution algorithm as implemented in the DEoptim (2.2-8) package in R. For each iteration of candidate parameter sets, the DMC simulated behavioural data — reaction time distributions split by response accuracy and congruency — which were compared with the observed data by minimising a  $G^2$  cost function. The  $G^2$  statistic was computed as

$$G^2 = 2 \sum_i O_i \ln \left( \frac{O_i}{E_i} \right), \quad (4)$$

where  $O_i$  and  $E_i$  are the observed and expected frequencies in bins defined by quantiles of the RT distribution, separately for correct and error responses in each congruency condition. For each parameter evaluation, the model generated 5 000 simulated trials using a time step of 1 millisecond. The differential evolution optimiser was run for a maximum of 1 000 iterations. After optimization, predictions from the best-fitting parameter set were generated using 20 000 simulated trials to obtain stable model predictions. Model output included the best-fitting parameter estimates, the associated cost value, and model comparison indices including AIC and BIC.

First, the aggregated dataset of all participants was fit to freely estimate all parameters globally. Subsequently, the time scale of the gamma distribution and shape of the starting point distribution were fixed to the mean values of the global fit. Following Vermeylen et al. (2025) these parameters were fixed to grand averages as they are not of interest to this study: The time scale signifies when interference peaks rather than the magnitude of interference and there is no theoretical reason to expect the starting point bias to differ because of the balanced experimental design.

These fixed values were used to compute the remaining parameters on per-participant behavioural data. The freely estimated parameters were (1) the controlled drift rate, (2) the decision boundary, (3) the amplitude of the automatic drift rate, (4) the mean of the non-decision time, and (5) the standard deviation of the non-decision time. Participant-level fits were estimated separately for each distance condition, but not for congruency, because congruency is explicitly modelled within the DMC framework rather than treated as a condition requiring separate parameter estimation. Specifically, congruency is implemented via a sign parameter  $c \in \{-1, 1\}$ , which modulates the

direction of the automatic activation component  $v_a(t)$ . This means that the same automatic activation can help performance on congruent trials but hinder it on incongruent trials.

Model fit was evaluated by calculating comparative goodness-of-fit measures that contrast the full model with a naïve restricted model that removes the amplitude and time scale parameters and accounts for conflict by specifying two different drift rates for congruent and incongruent trials. Specifically,  $\Delta G^2$  was computed as a formal likelihood ratio test to assess whether the improvement in fit was statistically significant, while  $\Delta AIC$  and  $\Delta BIC$  served as complexity-penalised model selection criteria to evaluate whether the added parameters were justified given the increase in model complexity. Additionally, simulated behavioural data was visually inspected and compared to the empirical data to further assess whether the model accurately captures the data.

To examine whether interfering information primarily modulated the amplitude parameter — indicating that the DMC appropriately captured the interference induced by the experimental factors — paired t-tests were computed on the participant-level parameter estimates comparing the distant and close conditions. The parameters tested were the controlled drift rate, decision boundary, mean non-decision time, and amplitude of the automatic activation. The tests produced estimates of the mean difference between conditions together with 95% confidence intervals. Effect sizes were quantified using Cohen's  $d$  for paired samples with 95% confidence intervals. Because multiple parameters were tested, p-values were adjusted using the Holm correction.

**Modelling Metacognitive Insight into Interference.** To formally answer the research question of whether individuals have insight into interfering information, we employed two distinct approaches. The first approach directly tested whether the DMC amplitude parameter predicted subjective interference ratings, and whether this relationship differed as a function of congruency. If participants are reporting interference not merely based on the observable properties of the trial — such as congruency or distance — but from an internal experience of the peak amplitude of automatic activation, their subjective ratings should reflect the direction and magnitude of that impulse. Because the automatic gamma pulse facilitates responding on congruent trials and interferes on incongruent trials, a participant tracking amplitude should report less interference as amplitude grows on congruent

trials and more interference as amplitude grows on incongruent trials. This opposing pattern across conditions was operationalised in the model as an interaction term between amplitude and congruency.

To test this, interference scores were averaged per participant and condition (distance x congruency). These condition-level means were then linked to the corresponding per-participant, per-distance amplitude estimates derived from the DMC fitting procedure. A LMM (referred to as model D) was fitted with the condition-level mean interference score as the outcome variable, with amplitude, congruency, and their interaction as fixed-effect predictors. A random intercept was included for each participant to account for individual differences in baseline interference ratings. Because participant-level DMC fits were estimated by distance rather than congruency, the same amplitude was linked to congruent and incongruent condition means. A significant interaction would indicate that the association between amplitude and subjective interference depends on trial type, consistent with participants having metacognitive sensitivity to the impulse driving behaviour rather than simply reporting trial congruency. Visual inspection of model diagnostics did not indicate any violations of assumptions, and the same transformation of estimates was applied as previously described to report EMMs.

The second approach employed a simulation-based procedure to further evaluate whether amplitude can be considered a direct driver of subjective interference. Using each participant's freely estimated DMC parameters, new behavioural data were generated trial-by-trial via forward simulation of the DMC. For each participant-condition combination, 1 000 trials were simulated. The amplitude was drawn from a normal distribution in each simulated trial, introducing realistic trial-to-trial variability in automatic activation. A simulated interference rating was then generated as a function of this trial-level amplitude, scaled by the congruency of the simulated trial. This operationalization encodes the theoretical assumption that, if participants have true insight into interference, their subjective ratings should scale directly with the amplitude of automatic activation.

To assess whether this theoretically motivated simulation recovers patterns consistent with the observed data, the four LMMs originally fitted to the behavioural data (models A, B, C and D) were re-estimated on the simulated dataset using identical model specifications, predictors, and random-

effects structures as the models based on empirical data. No violations were observed in any of the simulated model counterparts.

The resulting estimates and significance patterns from the simulated models were then compared visually and qualitatively to those from the observed models. Convergence between the two sets of models — particularly with respect to the direction and magnitude of effects — would suggest that the observed behavioural and metacognitive data are consistent with a generative process in which amplitude directly drives subjective interference. Divergence, by contrast, would indicate that additional or different mechanisms would be needed to account for the observed patterns. The simulation patterns will be graphically presented alongside the empirical results, before discussing their statistical counterparts.

**Table 1**

*Overview of Linear Mixed-Effects Models and Respective Specifications*

| Model | Specification                                                                           |
|-------|-----------------------------------------------------------------------------------------|
| A     | Log (Reaction Time) ~ Distance * Congruency + (1 + Congruency + Distance   Participant) |
| B     | Logit (Accuracy) ~ Distance * Congruency + (1 + Congruency + Distance   Participant)    |
| C     | Interference Score ~ Distance * Congruency + (1 + Congruency + Distance   Participant)  |
| D     | Interference Score ~ Amplitude * Congruency + (1   Participant)                         |

*Note. Model D is specified at condition-level data rather than trial-level data like models A-C.*

### 3 Results

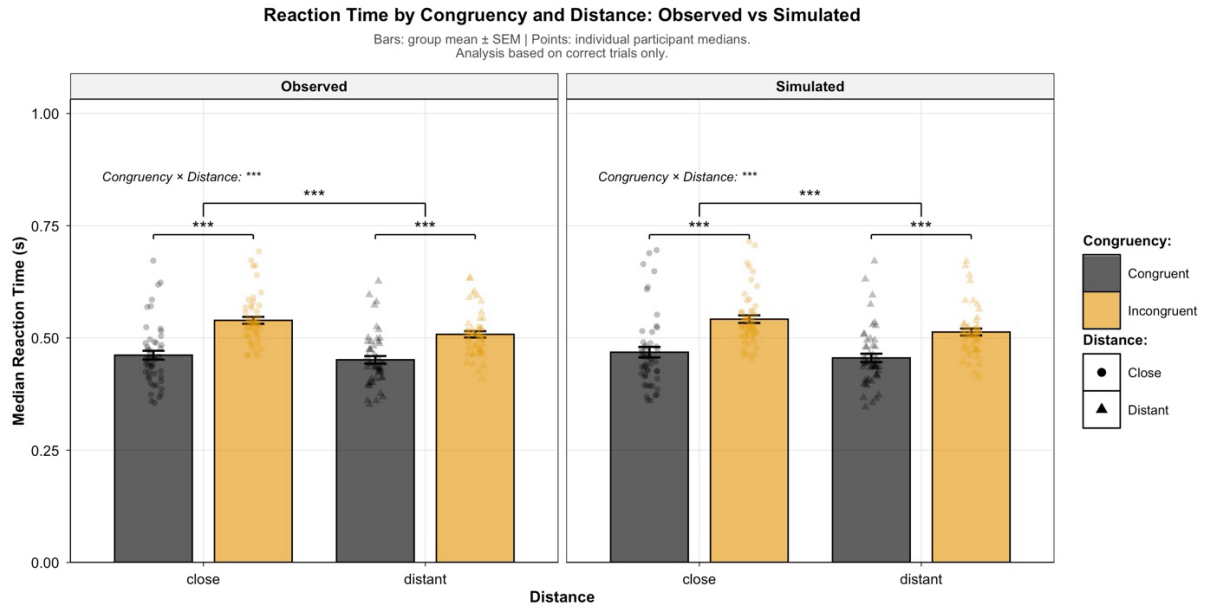
#### 3.1 Linear Mixed-Effects Modelling Results

To examine whether reaction times were influenced by our experimental manipulations (model A), log-transformed RTs were submitted to an LMM with congruency, distance, and their interaction as predictors. As shown in Figure 4, both congruency and distance had significant main effects on log-RTs,  $F(1, 48.9) = 214.462, p < .001, \eta^2_p = .81, 95\% \text{ CI } [.72, .87]$ , and  $F(1, 49.3) = 52.376, p < .001, \eta^2_p = .51, 95\% \text{ CI } [.32, .65]$ , respectively. A significant congruency x distance interaction was also observed, though its effect size was negligible ( $F(1, 18\ 394.6) = 14.365, p < .001, \eta^2_p < .001$ ).

Responses were fastest on congruent-distant trials (0.463 s, 95% CI [0.445, 0.481]) and slowest on incongruent-close trials (0.552 s, 95% CI [0.534, 0.571]), with congruent-close (0.479 s, 95% CI [0.458, 0.500]) and incongruent-distant (0.521 s, 95% CI [0.506, 0.536]) falling in between. Pairwise Tukey-corrected contrasts decomposed the interaction: Close stimuli produced slower RTs than distant stimuli for both congruent trials ( $z = 4.776, p < .001, d = 0.15, 95\% \text{ CI } [0.09, 0.22]$ ) and incongruent trials ( $z = 8.074, p < .001, d = 0.27, 95\% \text{ CI } [0.20, 0.33]$ ). Incongruent trials were slower than congruent trials in both the close condition ( $z = -15.036, p < .001, d = -0.64, 95\% \text{ CI } [-0.72, -0.56]$ ) and the distant condition ( $z = -12.426, p < .001, d = -0.53, 95\% \text{ CI } [-0.61, -0.44]$ ). As expected, the congruency effect was larger when flankers were presented close to the target than when they were presented farther away, indicating that spatial separation reduced the impact of conflicting information on response speed.

#### Figure 4

*Observed (left) and simulated (right) RTs align closely across conditions. Responses were fastest on congruent-distant trials and slowest on incongruent-close trials. Simulated data reproduced both the main effects and the interaction effect.*



Note. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

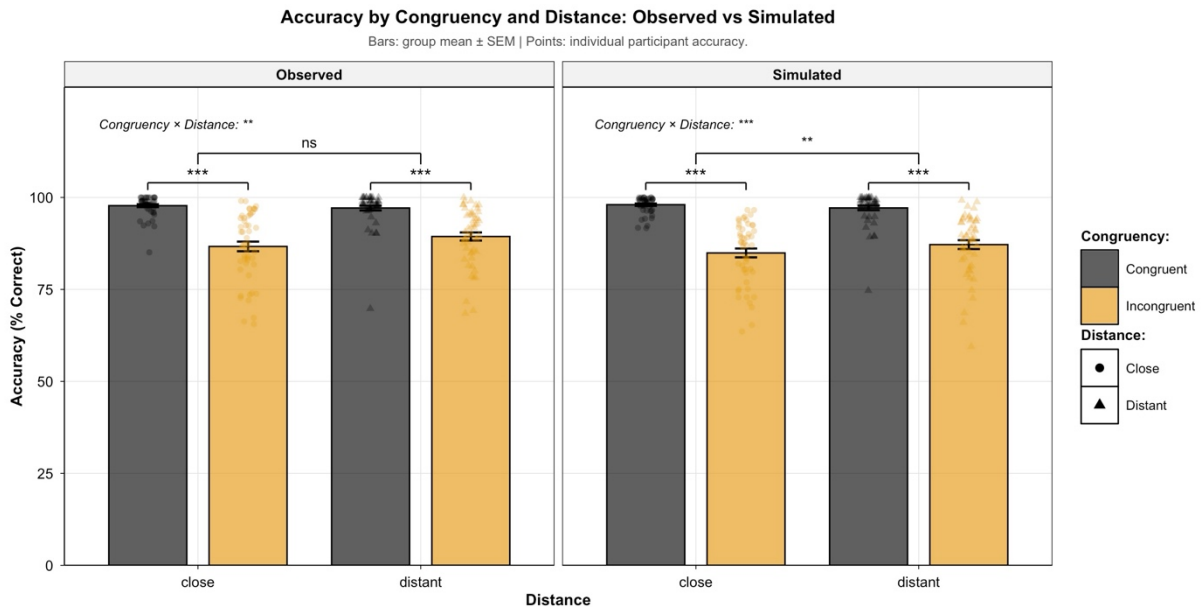
To assess whether accuracy was affected by the experimental manipulations (model B), trial-level responses were analysed using a logistic mixed-effects model with congruency, distance, and their interaction as predictors. A Type III Analysis of deviance revealed a significant main effect of congruency ( $\chi^2(1) = 129.203, p < .001$ ) and a significant congruency x distance interaction ( $\chi^2(1) = 9.487, p = .002$ ) as shown in Figure 5. Distance alone did not reach significance ( $\chi^2(1) = 1.388, p = .239$ ). The model explained 17.3% of the variance in accuracy based on fixed effects ( $R^2_m = .17$ ) and 33.4% when including both fixed and random effects ( $R^2_c = .33$ ).

Accuracy was near-ceiling for congruent trials (close: 98.5%, 95% CI [97.9, 98.9]; distant: 98.2%, 95% CI [97.5, 98.7]) but dropped substantially on incongruent trials — to 89.2% (95% CI [86.5, 91.4]) when close and 91.8% (95% CI [89.5, 93.6]) when distant.

Follow-up pairwise contrasts with Tukey correction showed that distance had no significant effect on accuracy for congruent trials ( $z = 1.178, p = .239$ ). For incongruent trials, accuracy was significantly lower for close than distant stimuli ( $z = -3.734, p < .001$ ), a statistically reliable but modest difference. Congruent trials resulted in higher accuracy than incongruent trials at both close ( $z = 11.367, p < .001$ ) and distant distances ( $z = 9.004, p < .001$ ). The interaction was driven by improved accuracy for incongruent trials at distant compared to close distances, reducing the congruency effect.

**Figure 5**

*Simulated accuracy (right) closely matched the empirical pattern (left): lower accuracy for incongruent trials, and an attenuation of the congruency effect at greater flanker distances.*



Note. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

To investigate whether subjective interference ratings varied as a function of our experimental manipulations (model C), ratings were entered into a LMM including congruency, distance, and their interaction as predictors. Congruency had a significant main effect on interference scores ( $F(1, 49.0) = 20.954, p < .001, \eta^2_p = .30, 95\% \text{ CI } [.11, .48]$ ), as did the congruency  $\times$  distance interaction ( $F(1, 18366.6) = 113.129, p < .001, \eta^2_p < .001$ ) as shown in Figure 6. Distance alone was not significant ( $F(1, 49.1) = 1.481, p = .230$ ).

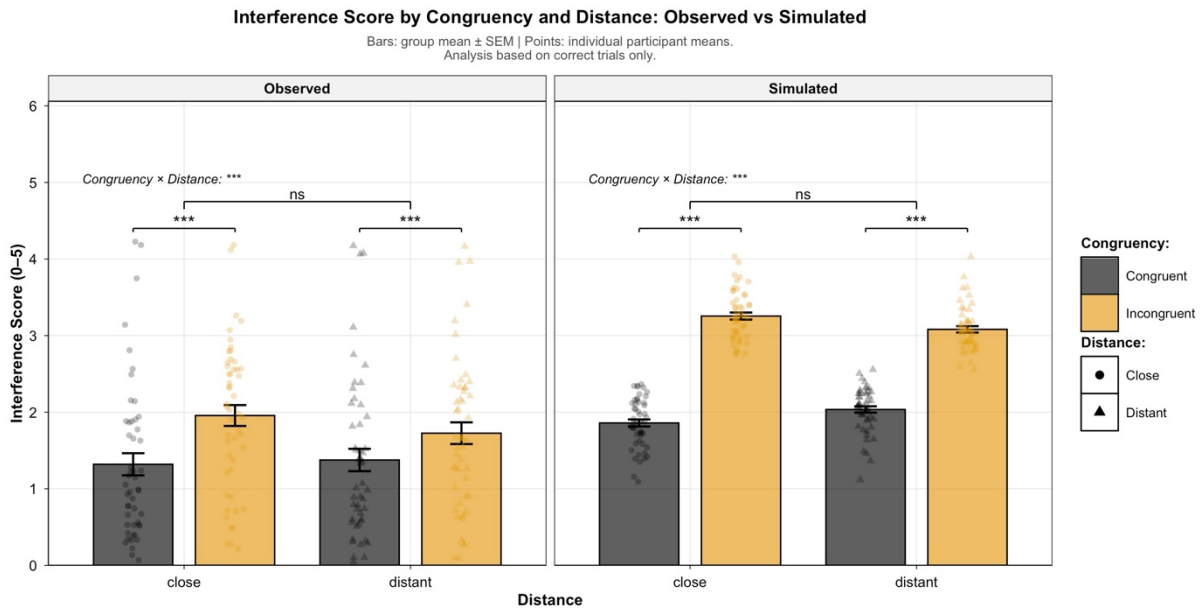
Incongruent trials attracted notably higher interference ratings (on a scale from 0-5) than congruent ones: 1.953 (95% CI [1.700, 2.207]) vs. 1.320 (95% CI [1.045, 1.595]) in the close condition, and 1.728 (95% CI [1.445, 2.012]) vs. 1.377 (95% CI [1.085, 1.669]) in the distant condition — a pattern mirroring the objective congruency manipulation.

Pairwise contrasts with Tukey correction revealed that for congruent trials, interference scores did not differ significantly between close and distant stimuli ( $z = -0.809, p = .418, d = -0.06, 95\% \text{ CI } [-0.22, 0.09]$ ). For incongruent trials, close stimuli produced higher interference scores than distant stimuli ( $z = 3.191, p = .001, d = 0.25, 95\% \text{ CI } [0.10, 0.40]$ ), a small-to-medium effect. Incongruent trials yielded higher interference scores than congruent trials at both close ( $z = -5.842, p < .001, d = -$

0.71, 95% CI [-0.94, -0.47]) and distant distances ( $z = -3.244$ ,  $p = .001$ ,  $d = -0.39$ , 95% CI [-0.63, -0.16]). The difference between congruent and incongruent trials was larger when stimuli were close and smaller when they were distant.

**Figure 6**

*Observed (left) and simulated (right) interference ratings showed similarity: incongruent trials received higher ratings, and the congruency effect was stronger for close stimuli.*



Note. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

### 3.2 Parameter Estimation

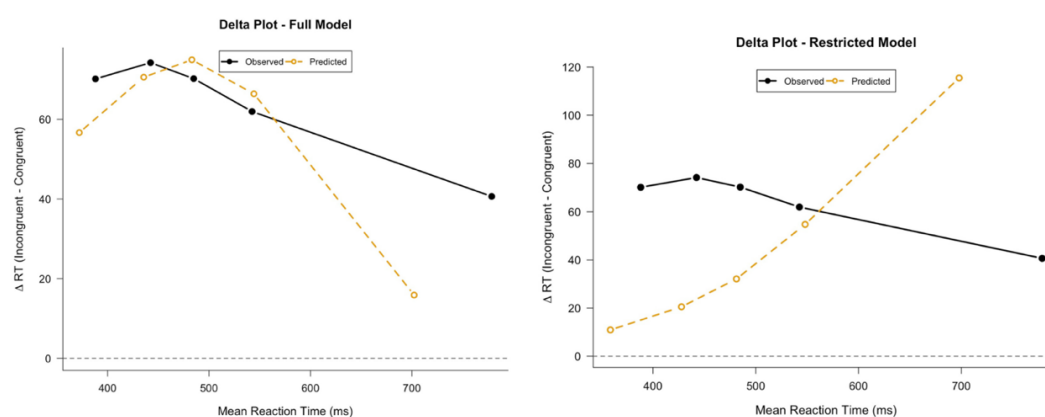
#### 3.2.1 Global Parameter Estimation

The DMC results indicate a good account of the empirical data. A likelihood ratio test and comparative fit indices confirm that the DMC (full model) fits significantly better than the restricted model — which only accounts for conflict by defining separate drift rates depending on trial congruency. This finding implies that the decomposed drift rate of the DMC, differentiating between a controlled drift rate and automatic activations captures interfering information well. The DMC fits significantly better than the restricted model ( $\Delta G^2(2) = 2302.600$ ,  $\Delta AIC = 2292.700$ ,  $\Delta BIC = 2300.600$ ) indicating that the conflict pulse parameters contribute significantly to explaining the empirical data. The aforementioned delta plots — showing how congruency effects evolve for different RTs — indicate that the full model captures the negative trajectory better than the restricted

model (as shown in Figure 7). Simulated RTs and accuracy rates of the DMC closely resemble the empirical data across both congruency and distance conditions, as shown in Figure 8. The time scale of automatic activation and the shape of the starting point distribution were fixed for participant-level estimations based on the global parameter estimation.

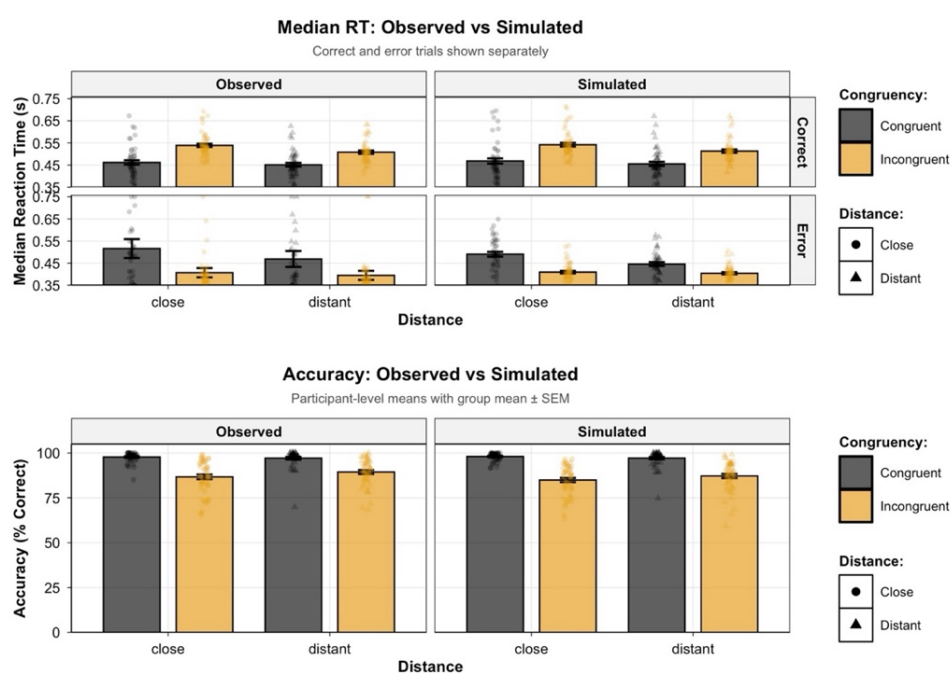
**Figure 7**

*Comparison of Delta Plots between the Full (left) and Restricted (right) Model. The predictions of the full model capture the empirical data better than the restricted model.*



**Figure 8**

*Behavioural Comparison between Empirical and Simulated Data.*



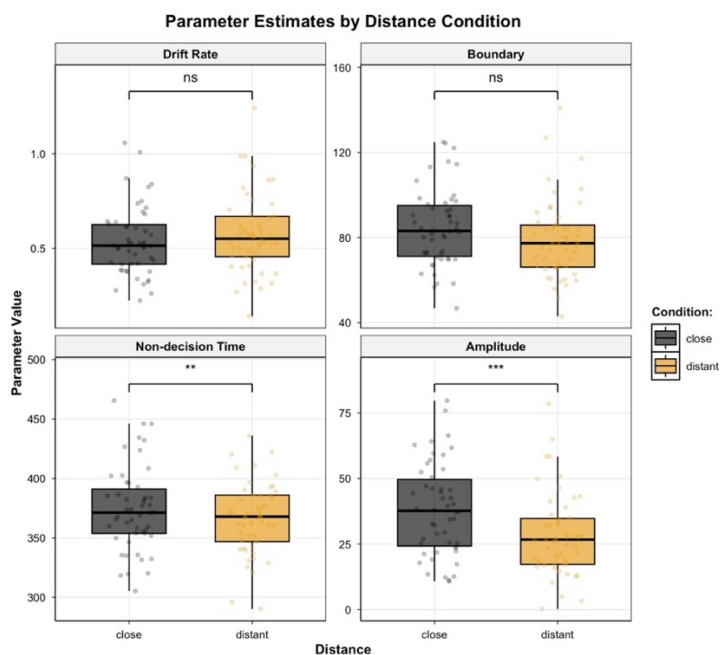
### 3.2.2 Participant-Level Parameter Estimation

After ensuring that the DMC captures the data well through global fit estimates, the model was fit separately to each participant's data for the close and distant conditions to examine whether DMC parameters differ as a function of spatial distance.

The paired t-tests indicate no significant differences between the close and distant conditions of the controlled drift rate after applying a Holm correction to the results to account for multiple comparisons ( $M_{\text{close}} = 0.54$ ,  $SD_{\text{close}} = 0.18$ ;  $M_{\text{distant}} = 0.58$ ,  $SD_{\text{distant}} = 0.21$ ),  $t(49) = 2.11$ ,  $p = .051$ ,  $d = 0.30$ , 95% CI [0.01, 0.58]. Similarly, the decision boundary did not significantly differ between conditions after correction for multiple testing ( $M_{\text{close}} = 84.6$ ,  $SD_{\text{close}} = 18.2$ ;  $M_{\text{distant}} = 79.0$ ,  $SD_{\text{distant}} = 18.7$ ;  $t(49) = -2.31$ ,  $p = .053$ ,  $d = -0.33$ , 95% CI [-0.61, -0.04]). In contrast, the mean non-decision time was significantly lower in distant trials ( $M_{\text{distant}} = 367$ ,  $SD_{\text{distant}} = 30.2$ ) compared to close trials ( $M_{\text{close}} = 374$ ,  $SD_{\text{close}} = 35.5$ ,  $t(49) = -3.41$ ,  $p = .004$ ,  $d = -0.48$ , 95% CI [-0.77, -0.19]). Finally, the amplitude of the automatic activation was significantly lower in the distant condition ( $M_{\text{distant}} = 28.4$ ,  $SD_{\text{distant}} = 15.9$ ) compared to the close condition ( $M_{\text{close}} = 38.2$ ,  $SD_{\text{close}} = 17.5$ ,  $t(49) = -5.41$ ,  $p < .001$ ,  $d = -0.77$ , 95% CI [-1.08, -0.45]).

**Figure 9**

*Overview of Paired T-tests between Conditions. Amplitude and non-decision time are significantly lower in distant conditions.*



### 3.3 Peak Amplitude and Subjective Interference

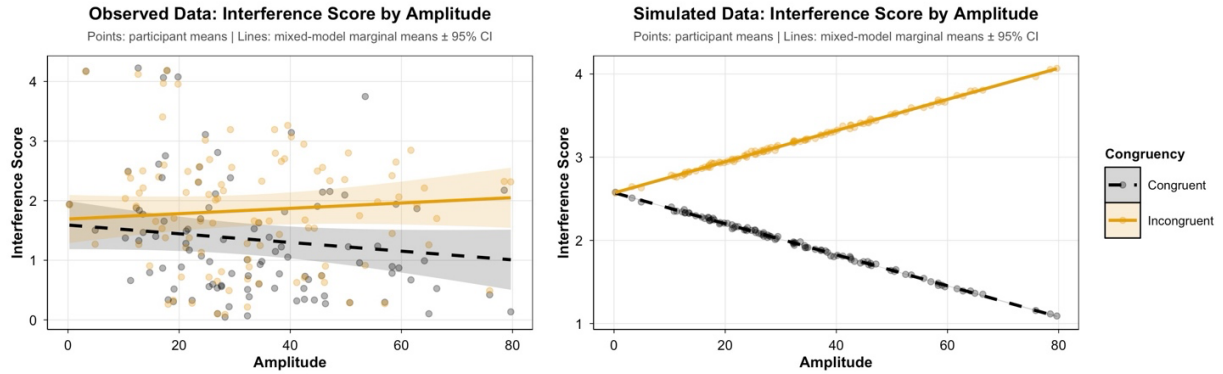
To test whether subjective interference reflects latent conflict processes (model D), interference ratings were modelled as a function of DMC amplitude, congruency, and their interaction using a linear mixed-effects model. Given the opposing directional influence of amplitude across congruency conditions, neither main effect was expected to reach significance. As hypothesized, neither main effect of amplitude and congruency reached significance,  $F(1, 195.93) = 0.12, p = .734$ , and  $F(1, 145.60) = 0.36, p = .550$ , respectively. The interaction between amplitude and congruency was significant ( $F(1, 145.60) = 7.00, p = .009, \eta^2_p = .05, 95\% \text{ CI } [.00, .13]$ ), indicating that participants' subjective experience of interference (on a scale of 0-5) tracked the congruency-specific influence of automatic activation on the decision process.

The divergence between conditions grew progressively with amplitude. At low amplitude (25<sup>th</sup> percentile, amplitude = 20.25), congruent and incongruent ratings were modestly separated (congruent: 1.443, 95% CI [1.156, 1.731] vs. incongruent: 1.782, 95% CI [1.495, 2.070]). At the median (amplitude = 29.35), the gap widened (congruent: 1.377, 95% CI [1.113, 1.641] vs. incongruent: 1.823, 95% CI [1.559, 2.087]). At high amplitude (75<sup>th</sup> percentile, amplitude = 44.47), congruent ratings had fallen to 1.266 (95% CI [0.986, 1.547]) while incongruent ratings rose to 1.891 (95% CI [1.610, 2.171]) — the two conditions clearly moving in opposite directions.

Amplitude trends revealed that the slope of amplitude on interference ratings differed significantly between conditions ( $b = -0.012, t(147) = -2.647, p = .009$ ), being negative for congruent trials ( $b = -0.007, 95\% \text{ CI } [-0.017, 0.002]$ ) and positive for incongruent trials ( $b = 0.004, 95\% \text{ CI } [-0.005, 0.014]$ ), though neither individual slope reached significance on its own. Tukey-adjusted pairwise contrasts at each amplitude percentile showed that incongruent trials produced higher interference scores than congruent trials at the 25<sup>th</sup> percentile ( $t(147) = -3.516, p < .001, d = -0.62$ ), the median ( $t(147) = -5.661, p < .001, d = -0.82$ ), and the 75<sup>th</sup> percentile ( $t(147) = -6.827, p < .001, d = -1.15$ ), indicating that the divergence between congruent and incongruent interference ratings became more pronounced as amplitude increased.

**Figure 10**

*As amplitude increased, congruent and incongruent interference ratings diverged in opposite directions. The interaction was significant in both the simulation and empirical data.*



### 3.4 Simulation Results

The simulation procedure was designed so that per-participant amplitude values derived from the DMC directly control the magnitude of the simulated interference rating, while trial congruency controlled the direction of this effect, determining whether interference was expressed toward the lower (no interference) or upper (very high interference) end of the rating scale. As previously mentioned, this simulation implements our hypothesis that interference ratings are a direct result of amplitude values of the DMC, signifying that individuals have metacognitive insight into interfering information during their decision process. Simulated behavioural data reflecting the empirical data were then calculated to establish to which degree these models share similarity.

To determine whether the simulated data reproduced the effects on response speed (simulation model A), log RTs were analysed using a LMM with congruency, distance, and their interaction as predictors. Both congruency and distance had significant main effects on log-RTs,  $F(1, 49) = 226.691$ ,  $p < .001$ ,  $\eta^2_p = .82$ , 95% CI [.73, .88], and  $F(1, 49) = 42.224$ ,  $p < .001$ ,  $\eta^2_p = .46$ , 95% CI [.26, .61], respectively. A significant interaction was also observed, though its effect size was again negligible ( $F(1, 91\ 711) = 107.205$ ,  $p < .001$ ,  $\eta^2_p < .001$ ).

Simulated condition EMMs closely mirrored the empirical data: congruent-distant trials were again fastest (0.464 s, 95% CI [0.447, 0.483]) and incongruent-close trials slowest (0.544 s, 95% CI

[0.527, 0.562]), with congruent-close (0.478 s, 95% CI [0.458, 0.500]) and incongruent-distant (0.514 s, 95% CI [0.499, 0.528]) reproducing the same ordering as the empirical data.

Pairwise contrasts with Tukey correction decomposed the interaction. Close stimuli produced slower RTs than distant stimuli for both congruent ( $z = 4.328, p < .001, d = 0.14, 95\% \text{ CI } [0.08, 0.21]$ ) and incongruent trials ( $z = 8.400, p < .001, d = 0.28, 95\% \text{ CI } [0.22, 0.35]$ ), with the distance effect roughly twice as large for incongruent trials. Incongruent trials were slower than congruent trials in both the close condition ( $z = -16.648, p < .001, d = -0.62, 95\% \text{ CI } [-0.70, -0.55]$ ) and the distant condition ( $z = -12.989, p < .001, d = -0.49, 95\% \text{ CI } [-0.56, -0.41]$ ). The congruency effect was medium-to-large at both distances but somewhat stronger for close stimuli, consistent with the interaction.

To evaluate whether the simulated data captured accuracy patterns (simulation model B), responses were analysed with a logistic mixed-effects model including congruency, distance, and their interaction as predictors. A Type III Analysis of deviance revealed a significant main effect of congruency,  $\chi^2(1) = 232.780, p < .001$ , a significant congruency x distance interaction,  $\chi^2(1) = 60.119, p < .001$ , and a significant main effect of distance ( $\chi^2(1) = 8.755, p = .003$ ). Unlike the empirical data, distance reached significance here, reflecting the increased statistical power of the larger simulated dataset. The model explained 25% of the variance in accuracy based on fixed effects ( $R^2_{\text{m}} = .25$ ) and 45% when including both fixed and random effects ( $R^2_{\text{c}} = .45$ ).

Simulated accuracy remained near-ceiling for congruent trials (close: 99.0%, 95% CI [98.6, 99.3]; distant: 98.7%, 95% CI [98.1, 99.1]), while accuracy decreased in incongruent trials: 86.8% (95% CI [84.4, 88.8]) when close and 89.5% (95% CI [87.1, 91.5]) when distant — reproducing the accuracy advantage for distant stimuli observed empirically.

Pairwise contrasts with Tukey correction showed that close stimuli produced lower accuracy than distant stimuli for congruent trials ( $z = 2.959, p = .003$ ) and for incongruent trials ( $z = -3.453, p < .001$ ). Congruent trials outperformed incongruent trials at both close ( $z = 15.257, p < .001$ ) and distant distances ( $z = 12.147, p < .001$ ). The interaction was driven by an attenuation of the congruency accuracy gap at distant compared to close distances.

To examine whether the simulated interference ratings reflected the experimental manipulations (simulation model C), ratings were analysed using a LMM with congruency, distance, and their interaction as predictors. Congruency had a significant main effect on interference ratings,  $F(1, 91\ 821) = 171.670, p < .001, \eta^2_p = .65, 95\% \text{ CI } [.65, .65]$ , as did the congruency x distance interaction,  $F(1, 91\ 798) = 3\ 455.7, p < .001, \eta^2_p = .04, 95\% \text{ CI } [.03, .04]$ . Distance alone was not significant, mirroring the empirical results ( $F(1, 91\ 795) = 0.664, p = .415, \eta^2_p < .001$ ).

Simulated ratings reproduced the qualitative pattern but with inflated magnitudes relative to the empirical data: congruent trials averaged 1.855 (close, 95% CI [1.846, 1.864]) and 2.030 (distant, 95% CI [2.021, 2.039]), while incongruent trials yielded 3.245 (close, 95% CI [3.236, 3.254]) and 3.075 (distant, 95% CI [3.065, 3.084]). Crucially, the cross-over across distance matched the empirical interaction.

Pairwise contrasts revealed that for congruent trials, interference ratings were lower in the close than the distant condition ( $z = -43.577, p < .001, d = -0.39, 95\% \text{ CI } [-0.41, -0.38]$ ), while the pattern reversed for incongruent trials ( $z = 39.728, p < .001, d = 0.38, 95\% \text{ CI } [0.37, 0.40]$ ). Incongruent trials produced higher interference ratings than congruent trials in both the close condition ( $z = -333.723, p < .001, d = -3.13, 95\% \text{ CI } [-3.16, -3.11]$ ) and the distant condition ( $z = -252.183, p < .001, d = -2.35, 95\% \text{ CI } [-2.38, -2.33]$ ), with the stronger congruency effect in the close condition driving the significant interaction. Note that the large effect sizes and  $z$ -values reflect the low residual noise characteristic of simulated data rather than practically larger effects than those observed empirically.

To assess whether the simulated ratings followed the hypothesized amplitude-driven pattern (simulation model D), interference scores were modelled as a function of amplitude, congruency, and their interaction. Like in the observed data, neither the main effect of amplitude nor the main effect of congruency reached significance,  $F(1, 80.04) = 0.40, p = .528$ , and  $F(1, 147.80) = 0.59, p = .445$ , respectively. The interaction between amplitude and congruency was significant ( $F(1, 147.80) = 94\ 838.02, p < .001, \eta^2_p = 1.00, 95\% \text{ CI } [1.00, 1.00]$ ).

Amplitude trends decomposed by condition confirmed that the slope of amplitude on interference ratings was equal in magnitude but opposite in direction across conditions: negative for congruent trials ( $b = -0.019, 95\% \text{ CI } [-0.019, -0.019]$ ) and positive for incongruent trials ( $b = 0.019,$

95% CI [0.019, 0.019]), and these slopes differed significantly from one another ( $b = -0.037$ ,  $t(148) = -307.96$ ,  $p < .001$ ). The congruency contrast also grew systematically with amplitude: at the 25<sup>th</sup> percentile (amplitude = 20.25), interference scores were lower for congruent (2.195, 95% CI [2.191, 2.199]) than incongruent trials (2.949, 95% CI [2.946, 2.953],  $d = -50.9$ ,  $p < .001$ ); at the median (amplitude = 29.35), the difference increased (congruent: 2.025, 95% CI [2.022, 2.028]; incongruent: 3.120, 95% CI [3.117, 3.123];  $d = -73.9$ ,  $p < .001$ ); and at the 75<sup>th</sup> percentile (amplitude = 44.47), the divergence was largest (congruent: 1.743, 95% CI [1.739, 1.747]; incongruent: 3.404, 95% CI [3.400, 3.407],  $d = -112.1$ ,  $p < .001$ ). This pattern confirms that the simulation is functioning as intended: the model successfully recovers the hypothesized interaction.

## 4 Discussion

This study set out to examine whether individuals have metacognitive insight into the degree to which task-irrelevant information interferes with decision-making. The DMC was fitted to behavioural data to extract latent parameters, and these parameters were linked to subjective reports. Behavioural and computational results provided insight into the relationship between amplitude and the subjective experience of interference during task performance.

The models predicting RT and accuracy showed that stimulus congruency and distance modulate RTs and accuracy scores, allowing us to understand whether interference takes place behaviourally. The RT results showed strong evidence for congruency effects. Participants reacted slower during incongruent trials compared to congruent trials, indicating that conflicting information impedes processing efficiency — consistent with the well-established congruency effect: responses are slower and less accurate when flankers point in the opposite direction to the target, reflecting interference from automatically activated competing response tendencies (Botvinick et al., 2001; Eriksen & Eriksen, 1974). A smaller but still significant effect of distance showed that distant distractors had less influence compared to close distractors, suggesting visual separation reduces interference. The significant interaction reflected a smaller congruency effect when flankers were presented at greater distances. This suggests that spatial separation attenuates the degree to which task-irrelevant flankers interfere with target processing.

As expected, participants were less accurate on incongruent trials compared to congruent trials. This shows that conflict not only slows down response times but also affects accuracy. Distractor distance alone did not significantly affect accuracy scores, but a significant interaction emerged: conflict effects were attenuated when stimuli were more distant. Incongruent trials were less error-prone in visually separated scenes compared to close ones. This suggests spatial separation can buffer against conflict-related performance decreases. While spatial separation might not improve accuracy on its own, it plays a key role in conflicting scenes by reducing the interference of incompatible information. The RT and accuracy findings also closely resemble the findings of Fischer et al. (2018), where congruency effects and distance effects reflect the modulation of interference strength, with reduced interference leading to improved performance.

Model C extended the findings by showing whether subjective interference ratings also track congruency and distance. Interference ratings were significantly higher for incongruent compared to congruent trials, showing that participants were able to make metacognitive judgements about their perceived levels of interference as expected. This aligns with Maniscalco and colleagues (2012), who showed using a modified flanker task that subjective perception of distracting flankers heightened interference effects during the task. Participants did not perceive lower conflict for distant stimuli when compared to close stimuli. The modest interaction effect showed that distance shaped subjective ratings in high-conflict trials, similar to accuracy measures. Incongruent close trials were rated as more conflicting than incongruent distant trials, whereas congruent trials showed no such difference. Subjective conflict ratings were modulated by distance only when interference was introduced in the trial. These findings imply that, specifically on incongruent distant trials, participants not only perform better but also feel that there is less conflict. This pattern is consistent with Questienne and colleagues (2021), who argue that subjective awareness plays a functional role in cognitive control mechanisms, with multiple factors — like congruency and spatial separation — shaping subjective experience of interference.

A consistent pattern is visible across these three outcome variables behaviourally: congruency consistently produced faster RTs, higher accuracy and lower ratings of interference, showcasing the already known conflict effects in similar paradigms. The role of distance is more complex: While distant trials lead to faster RTs, it does not affect accuracy or interference ratings. The influence of distance only appears in already conflicting situations. Both accuracy and interference ratings demonstrate the effect distance has on incongruent information, indicating that spatial distance reduces the cognitive effort imposed by congruency effects. Spatial separation alleviates conflict-related processing costs across objective and subjective measures alike. That subjective ratings mirror objective performance in this way supports the notion that congruency effects and trial-by-trial adaptations in conflict tasks have a subjective counterpart — performance is not only shaped by objective conflict, but also by how strongly that conflict is felt (Desender et al., 2017).

By fitting the DMC to behavioural data, the RT and accuracy patterns were decomposed into latent cognitive parameters. The simulated data showed close resemblance to the empirical data and

the comparative fit indices showed substantial improvement, indicating that the DMC captured the data well compared to a more restricted model that only accounts for conflict by defining separate drift rates for congruent and incongruent trials. The paired t-tests between close and distant trials confirm that the amplitude parameter captures interference and corresponds to the experimental manipulation as hypothesized. We were able to replicate the notion that peak amplitude is a meaningful index of interference. Surprisingly, the non-decision time also differed significantly based on distance: participants spent a longer time encoding perceptual information in the close trials. As non-decision time captures the time taken for perceptual processing before evidence accumulation begins, a straightforward interpretation is that the perceptual system may require more time to parse the target from its immediate surround when visual stimuli are closer to each other. Close distractors are more likely to capture attention automatically due to spatial proximity effects, meaning the perceptual system may initially process the wrong location before reorienting to the target.

The prior results (congruency and distance effects on RT, accuracy, and subjective ratings, and the superior fit of the DMC) largely replicate and extend established findings (Fischer et al., 2018; Ulrich et al., 2015). The novel contribution of the present work lies in the direct test of whether the DMC's peak amplitude parameter predicts subjective interference reports in a condition-specific way. This tests not merely whether people experience more conflict in incongruent trials, but whether the magnitude of automatic capture as quantified by the peak amplitude parameter is reflected in metacognitive reports. The condition-level mixed-effects model linked participants' amplitude estimate to condition-level interference ratings and examined to which degree amplitude predicts how much interference someone reports feeling, further specifying this question by assessing whether this relationship differs between congruent and incongruent trials.

The insignificant main effect of amplitude shows that participants did not uniformly report more interference regardless of whether the automatic impulse pulled them towards the correct boundary. Considering that the automatic gamma pulse of the DMC pushes evidence towards the correct boundary in congruent trials and pulls evidence away during incongruent trials, this non-significant result was expected. Similarly, the insignificant main effect of congruency indicates that participants did not solely report incongruent trials as more interfering than congruent trials.

On the contrary, the significant interaction effect shows that as amplitude increases, interference reports decrease slightly on congruent trials and increase on incongruent trials, reflecting the role of the gamma pulse in each condition. This contrast grows with the size of the amplitude, as is evident by assessing differences at amplitude percentiles. Participants are not just distinguishing by congruency but are sensitive to the magnitude of the automatic gamma pulse.

This interpretation goes beyond the account offered by Desender and colleagues (2017), who proposed that subjective difficulty reflects the integration of multiple observable cues such as reaction time and response conflict. The present findings suggest that participants have access to the latent state of the automatic activation process itself. Interestingly, this would also explain why ratings appear to depend on multiple superficial cues in the Desender et al. (2017) framework: because amplitude is itself related to both RT and congruency, tracking amplitude would produce data that looks like integration of multiple cues, while actually reflecting a single conflict parameter. The idea that the latent amplitude parameter itself predicts subjective interference ratings in a condition-specific and graded manner shows that participants integrate internal conflict signals into their metacognitive reports.

Going beyond the regression approach, we also evaluated whether the observed metacognitive data are consistent with a generative model in which peak amplitude is the direct driver of subjective interference. This simulation-based approach complemented the regression approach by asking a different question: not merely whether amplitude predicts ratings, but whether a model that assumes amplitude drives ratings can reproduce the full pattern of empirical data across all outcome variables. Comparing the LMMs based on empirical data to simulated counterparts provided a more nuanced perspective next to model D. Across the LMMs, the simulated data reproduced the core qualitative patterns of the empirical data remarkably well. All significant effects in the empirical models were also significant in the simulated models. Crucially, all insignificant effects were also non-significant in the simulated counterparts, with one exception. Effect sizes were generally similar between both approaches, indicating that the simulations preserved the size of behavioural effects.

The most notable differences emerged in Models C and D, and they are mostly attributable to the simulation technique implemented. In model C, the empirical model showed substantial random

slope variance, but the simulated data could not support a complex random-effects structure — a random-intercepts-only model was required. Moreover, effect sizes in simulated model C were substantially inflated. In model D, the empirical model yielded a small interaction ( $\eta^2_p = .05$ ), whereas the simulated model produced a perfect effect ( $\eta^2_p = 1.00$ ) and individual slopes that were highly significant on their own — neither of which was true empirically. These discrepancies are expected artefacts of simulation: extremely low residual variance and large sample size produce artificially inflated test statistics, but they do not indicate a failure of the simulation. On the contrary, the fact that the simulated data — generated under the explicit assumption that subjective interference scales with trial-level amplitude — successfully recovered the identical qualitative interaction pattern provides strong convergent evidence that the empirical findings are consistent with a generative process in which amplitude directly drives metacognitive reports.

Prior work has suggested that subjective feelings of interference are constructed post-hoc from observable performance cues such as RTs and congruency (Desender et al., 2017; Questienne et al., 2021). The present findings challenge this view. The directionally specific amplitude-by-congruency interaction where the same increase in amplitude drives ratings in opposite directions depending on trial type, is inconsistent with a post-hoc readout of congruency or RT and instead suggests that participants have access to the magnitude of automatic activation signal itself. The present findings therefore offer a more parsimonious explanation. This reframing has direct consequences for how cognitive control is understood. A person with access to amplitude can calibrate suppression in proportion to the actual strength of the automatic impulse, rather than responding uniformly to the mere presence of conflict — a more efficient and targeted control strategy. Beyond its theoretical relevance, this finding also suggests that subjective reports are not epiphenomenal but carry information not fully captured by behavioural measures alone, opening the possibility of using them as additional constraints when fitting computational models. More broadly, populations characterised by impaired cognitive control may differ not only in how strongly interference affects their decisions, but in whether they have sufficient metacognitive access to that interference to regulate their behaviour adaptively.

## 4.1 Limitations and Directions for Future Research

Although the present findings provide evidence that individuals have metacognitive insight into interfering information during decision-making, some limitations should be considered. The first concerns the specification of model D, which was estimated at the level of conditions rather than trials. This followed from the structure and constraints of the modelling pipeline: the DMC yielded amplitude estimates at the level of each participant and distance condition, rather than an amplitude value by trial. The LMM therefore linked condition-level mean interference ratings to condition-level amplitude estimates. While this approach is justified and sufficient to test whether subjective reports track objective conflict at an aggregated level, it does not allow for a direct trial-by-trial test of whether momentary fluctuations in automatic activation predict subjective interference. The present analysis supports the claim that subjective interference covaries with conflict strength across conditions, but does not yet provide a higher resolution account of metacognitive insight. Future work would benefit from extending the DMC model so that a latent amplitude value can be estimated for each individual trial. Such a model could provide a more fine-grained account of metacognitive insight and clarify more precisely how subjective interference emerges during the decision process itself.

A second limitation entails that subjective awareness is likely shaped by multiple cues rather than a single latent parameter. This notion is supported when comparing the empirical and simulated models of the regression of congruency and amplitude on interference ratings: the empirical data resembles a noisy read-out of the simulated relationship between amplitude and interference, indicating there could be more factors related to interference ratings than amplitude only (as is seen in Figure 10). This idea is in line with previous work showing that metacognitive judgements of difficulty and conflict can depend on several sources of information, such as conflict itself, response dynamics, and other task-related signals (Desender et al., 2017). The simulation procedure used in this thesis deliberately imposed the strong assumption that subjective interference scaled directly with amplitude, and the fact that the simulated data qualitatively resembled the empirical pattern provides important convergent support for that account. Nevertheless, this assumption likely simplifies the true complexity of metacognitive judgements. Future studies could therefore compare richer and nested models in which subjective interference is jointly predicted by, next to trial-level conflict parameters,

other latent parameters of the decision process in order to determine which combination of factors best explains subjective experience.

A third limitation is that the present study relied on a single conflict paradigm, namely an Eriksen-flanker task with a spatial distance manipulation. This makes the current conclusions theoretically strong within the flanker framework, but it remains unclear to what extent the observed relationship between latent conflict parameters and subjective interference generalises to other conflict tasks. The present findings should not yet be interpreted as demonstrating a fully domain-general mechanism of metacognitive insight into conflict.

A final concern is the conceptual distinction between interference and difficulty at the level of subjective reports. Although participants were instructed to rate interference rather than task difficulty, it is possible that they conflated the two, relying on cues such as response fluency or uncertainty rather than isolating the specific influence of task-irrelevant information. If this were the case, any measure correlated with task effort should predict subjective ratings and not specifically the peak amplitude in interaction with congruency. The observed interaction where the direction of the amplitude effect reverses between congruent and incongruent trials, is inconsistent with a general difficulty judgement — which would predict a uniform positive relationship with amplitude. This dissociation suggests that subjective ratings were in fact sensitive to interference specifically rather than difficulty alone, aligning with our hypothesis.

Future research should aim to replicate these findings to establish the robustness of the relationship between amplitude and interference. The validity of these findings may also benefit from extending these analyses to other classic conflict tasks such as the Stroop and Simon tasks. This will determine whether the same parameter mapping emerges across different modalities of interference and differentiates a general feature of cognitive control from task-specific findings. Finally, because diffusion modelling is often linked to neural measures in the literature, integrating neuroimaging methods could help identify neural correlates of amplitude, conflict monitoring and metacognitive evaluation. This would provide a stronger and holistic account for how people detect and experience interference during decision-making.

## 5 Conclusion

This thesis set out to examine whether people have metacognitive insight into the extent to which task-irrelevant information interferes with speeded decision-making. This question was addressed by leveraging diffusion modelling techniques and simulation analyses applied to empirical data of a modified Eriksen-flanker task. By combining behavioural, computational and metacognitive perspectives, these findings provide converging evidence that interfering information is not only expressed in response times and accuracy but is also reflected in how participants subjectively experience the decision process.

Congruency effects were observed across outcome measures, while greater spatial distance did not uniformly improve performance but rather reduced the cost of conflict. These findings suggest that spatial separation mainly matters when distractors compete with task-relevant information, attenuating the disruptive effect of that competition rather than producing a general facilitation across all trial types. The DMC not only fit the data well but also isolated a theoretically meaningful parameter that tracks the effect of the experimental manipulation. Participant-level parameter estimates showed that flanker distance primarily reduced the amplitude of the automatic activation and, to a lesser extent, non-decision time. Within the present task, this pattern implies that spatial distance mainly altered the strength of task-irrelevant capture compared to other aspects of the decision process.

The main contribution of this work lies in the link between peak amplitude as a latent interference parameter and subjective experience. Our findings support the notion that the relationship between amplitude and interference experiences depends on stimulus congruency. The automatic activation encoded by the DMC has opposite meanings across congruency conditions. The significant amplitude-congruency interaction therefore suggests that participants' subjective ratings varied in a way consistent with the direction and size of the automatic impulse imposed on the decision process. Subjective reports of interference contained metacognitive information about the latent mechanisms underlying the decision process.

The simulation analysis further strengthened this conclusion by asking what the data should look like if subjective interference was directly driven by variation in amplitude. Across the mixed-effects models on the simulated data, the qualitative structure of the empirical data was recovered

remarkably well. The same core pattern emerged for reaction times, accuracy, and interference ratings, and the critical amplitude-by-congruency interaction was recovered as hypothesized.

Overall, this thesis provides evidence to support the idea that metacognitive insight into interfering information is present. Subjective awareness of interference does not appear to simply reflect congruency or a uniform increase in perceived difficulty. Instead, it tracks the way in which task-irrelevant information shapes the decision process. By bringing together behavioural measures, subjective reports, diffusion modelling, and simulation, the present work offers a more integrated account of conflict monitoring in decision-making. The findings support the view that cognitive control and metacognitive monitoring are closely intertwined: Individuals have insight into the extent to which irrelevant information is biasing their decision and consequently adapt to conflict behaviourally. Beyond its theoretical contributions, the integration of subjective reports and computational modelling offer a richer account of decision-processes that may prove valuable in understanding how metacognitive awareness — or lack thereof — contributes to maladaptive behaviour in clinical populations characterised by impaired cognitive control.

This thesis therefore contributes to a more holistic understanding of adaptive decision-making by showing that subjective experiences of interference can meaningfully inform the study of cognitive control rather than standing apart from it.

## 6 References

- Bognar, M., Szekely, Z., Varga, M. A., Nagy, K., Spinelli, G., Hartanto, A., Majeed, N. M., Chen, N. R., Gyurkovics, M., & Aczel, B. (2024). Cognitive control adjustments are dependent on the level of conflict. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-53090-4>
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624–652. <https://doi.org/10.1037//0033-295x.108.3.624>
- Desender, K., Van Opstal, F., & Van Den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological Science*, 25(3), 675–683. <https://doi.org/10.1177/0956797613511468>
- Desender, K., Van Opstal, F., & Van Den Bussche, E. (2017). Subjective experience of difficulty depends on multiple cues. *Scientific Reports*, 7(1). <https://doi.org/10.1038/srep44222>
- Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, 29(5), 761–778. <https://doi.org/10.1177/0956797617744771>
- Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 380–390. <https://doi.org/10.3758/cabn.7.4.380>
- 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>
- Fischer, A. G., Nigbur, R., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2018). Cortical beta power reflects decision dynamics and uncovers multiple facets of post-error adaptation. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-07456-8>
- Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, 121(4), 480–506. <https://doi.org/10.1037//0096-3445.121.4.480>

- Huang, Y.-T., Georgiev, D., Foltynie, T., Limousin, P., Speekenbrink, M., & Jahanshahi, M. (2015). Different effects of dopaminergic medication on perceptual decision-making in Parkinson's disease as a function of task difficulty and speed-accuracy instructions. *Neuropsychologia*, 75, 577–587. <https://doi.org/10.1016/j.neuropsychologia.2015.07.012>
- Huang-Pollock, C., Ratcliff, R., McKoon, G., Shapiro, Z., Weigard, A., & Galloway-Long, H. (2016). Using the diffusion model to explain cognitive deficits in attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology*, 45(1), 57–68. <https://doi.org/10.1007/s10802-016-0151-y>
- 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>
- Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *The Journal of Neuroscience*, 33(50), 19434–19441. <https://doi.org/10.1523/jneurosci.3355-13.2013>
- Maniscalco, B., Bang, J. W., Iravani, L., Camps-Febres, F., & Lau, H. (2012). Does response interference depend on the subjective visibility of flanker distractors? *Attention, Perception, & Psychophysics*, 74(5), 841–851. <https://doi.org/10.3758/s13414-012-0291-2>
- Moustafa, A. A., Kéri, S., Somlai, Z., Balsdon, T., Frydecka, D., Misiak, B., & White, C. (2015). Drift Diffusion Model of reward and punishment learning in schizophrenia: Modeling and experimental data. *Behavioural Brain Research*, 291, 147–154. <https://doi.org/10.1016/j.bbr.2015.05.024>
- Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-018-01193-y>
- Questienne, L., van Dijck, J.-P., & Gevers, W. (2021). The role of subjective experiences in conflict tasks: A Review. *Psychologica Belgica*, 61(1), 46–62. <https://doi.org/10.5334/pb.508>
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281. <https://doi.org/10.1016/j.tics.2016.01.007>

- 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>
- 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>
- 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>
- 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).  
<https://doi.org/10.1038/s44271-025-00282-x>
- White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the Flanker Task: Discrete versus gradual attentional selection. *Cognitive Psychology*, 63(4), 210–238.  
<https://doi.org/10.1016/j.cogpsych.2011.08.001>