

ALAN VOODLA

Does it feel right?
Towards an affective perspective
on metacognitive confidence



KU LEUVEN



KU LEUVEN

FACTULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES



UNIVERSITY
OF TARTU

DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS

67

ALAN VOODLA

Does it feel right?
Towards an affective perspective on
metacognitive confidence

Doctoral thesis offered to obtain the joint degree of
Doctor of Psychology (PhD) from KU Leuven and
Doctor of Philosophy (Psychology) from University of Tartu

Supervisor(s):

prof. dr Andero Uusberg (UT),
prof.dr Kobe Desender (KUL)



UNIVERSITY OF TARTU
Press

Institute of Psychology, University of Tartu, Estonia
Faculty of Psychology and Educational Sciences, KU Leuven, Belgium

The dissertation is submitted for defence for a joint doctoral degree in Psychology: Doctor of Philosophy, University of Tartu (home institution), March 30, 2026, and Doctor of Psychology, KU Leuven (host institution), March 20, 2026.

Supervisors: Andero Uusberg, PhD, Professor
University of Tartu, Estonia
Kobe Desender, PhD, Associate Professor
KU Leuven, Belgium

Opponent: Andrey Chetverikov, PhD, Associate Professor
University of Bergen, Norway

Examination Committee members:
Agnes Moors, PhD, Professor, KU Leuven, Belgium
Jaan Aru, PhD, Associate Professor, University of Tartu, Estonia
Jüri Allik, PhD, Professor, University of Tartu, Estonia

Commencement: May 6, 2026 in Näituse 2, Tartu

Publication of this thesis is granted by the Institute of Psychology, University of Tartu and by the Faculty of Psychology and Educational Sciences, KU Leuven, Belgium.

ISSN 1024-3291 (print)
ISBN 978-9908-57-177-5 (print)
ISSN 2806-2531 (pdf)
ISBN 978-9908-57-178-2 (pdf)

Copyright: Alan Voodla, 2026

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS	6
1. INTRODUCTION.....	7
2. CONFIDENCE AND AFFECT AS METACOGNITIVE SIGNALS	11
Study I.....	14
3. MECHANISMS OF METACOGNITIVE AFFECT	17
Study II.....	18
Study III	19
4. AffectDDM – A COMPUTATIONAL FRAMEWORK OF METACOGNITIVE AFFECT	23
5. NEURAL CORRELATES OF METACOGNITIVE SIGNALS.....	28
Study IV	29
6. SUMMARY AND CONCLUSIONS.....	32
ACKNOWLEDGEMENTS	34
REFERENCES.....	35
KOKKUVÕTE.....	41
SAMENVATTING	42
SUMMARY	43
PUBLICATIONS	45
CURRICULUM VITAE	125
ELULOOKIRJELDUS.....	126

LIST OF ORIGINAL PUBLICATIONS

Study I – Voodla, A., Uusberg, A., & Desender, K. (2025). Metacognitive confidence and affect – two sides of the same coin?. *Cognition and Emotion*, 1–18.

Study II – Voodla, A., Uusberg, A., & Desender, K. (2024). Affective valence does not reflect progress prediction errors in perceptual decisions. *Cognitive, Affective, & Behavioral Neuroscience*, 24(1), 60–71.

Study III – Vrantsidis, T., Voodla, A., & Sabbagh, K. (2025). Multilevel Reward Prediction Errors, Not Expectations or Outcomes, Drive Emotional Valence. (pre-print)

Study IV – Voodla, A., & Uusberg, A. (2021). Do performance-monitoring related cortical potentials mediate fluency and difficulty effects on decision confidence?. *Neuropsychologia*, 155, 107822.

The author of the thesis contributed to studies I, II and IV as the main author by leading the development of the theory, design and methods, collecting and analyzing the data and writing up the manuscript. For study III the author of the thesis was in a supportive role in all the phases of the research process.

1. INTRODUCTION

“Doubt is an uneasy and dissatisfied state from which we struggle to free ourselves and pass into the state of belief; while the latter is a calm and satisfactory state which we do not wish to avoid, or to change to a belief in anything else.”

– Charles Sanders Peirce, *The fixation of belief* (1877)

This thesis traces back to my first encounter with Peirce’s *The Fixation of Belief* (1877) during my Bachelor studies, where the passage above sparked my interest in how we form, sustain, and sometimes revise our beliefs and decisions. When I turned to the literature of cognitive psychology, I learned about the concepts and theories of metacognition and decision confidence, where the emotional and motivational aspects highlighted in the quote by Peirce were rarely discussed or experimentally studied. This tension laid the foundation for the central question of this thesis: what is the role of affect in metacognition?

Humans are not only thinkers, but also thinkers about their thinking – a capacity known as metacognition. This ability to monitor and regulate our mental states, such as beliefs, decisions, and emotions, is crucial for learning, problem-solving, and flexible behavior (Efklides, 2006; Flavell, 1979; Fleming, 2024; Gross, 2015; Thompson et al., 2011). A core aspect of metacognition is confidence – our internal evaluation of the accuracy of decisions and beliefs (Grimaldi et al., 2015; Pouget et al., 2016; Sharot et al., 2023; Yeung & Summerfield, 2012). Confidence guides our behavior by telling us whether to stick with a decision or explore alternatives. This can occur in decisions about how to perceive the world (i.e. perceptual decisions) as well as in decisions about how to act upon it (i.e. value-based decisions). In both cases, confidence functions like a thermostat for the mind. Just as a thermostat maintains room temperature by detecting deviations and calling for adjustments, confidence calibrates to what extent we commit to our decisions or act upon beliefs. When confidence is well-calibrated, our decisions appropriately reflect the uncertainty and complexity of the world. When confidence is systematically too low or too high, maladaptive patterns such as overthinking, anxiety, impulsivity, or resistance to feedback can emerge (Hoven et al., 2019; Rouault et al., 2018).

The focus of this thesis is on the affective aspects of confidence. Confidence is often conceptualized as a cognitive probability estimate – the subjective likelihood that one is accurate in their decisions (Kepecs et al., 2008; Pouget et al., 2016; Sanders et al., 2016). Yet in everyday experience, confidence also feels a certain way. The sense of being certain is accompanied by ease or relief, while the sense of doubt carries a feeling of tension or discomfort (Ackerman & Thompson, 2017; Arango-Muñoz, 2014; Efklides, 2006; Koriat, 2000). From an affective science perspective such feelings can be understood as instances of core affect – transient states varying along the dimensions of valence (pleasant–

unpleasant) and arousal (activation–deactivation) (Barrett, 2017; Russell, 2003). Confidence thus qualifies as a metacognitive affective state: a feeling about one’s decision accuracy, akin to feelings of insight, knowing or difficulty (Loev, 2022; Velasco & Loev, 2025). What is yet to be empirically investigated, is how metacognitive affect emerges and whether it tracks decision accuracy similarly to judgments of confidence.

Various theories propose different ways of thinking about how confidence and affect might be related (discussed more thoroughly in Chapters 2 and 3). In perceptual metacognition literature that focuses on computational modelling of confidence, the affective aspects of metacognition are largely ignored (Kepecs et al., 2008; Pouget et al., 2016; Yeung & Summerfield, 2012). Although good candidates for computational and neural mechanisms underlying confidence evaluations have been proposed (Fleming & Daw, 2017; Fleming & Dolan, 2012; Meyniel et al., 2015; Shekhar & Rahnev, 2024), these approaches typically focus on the accuracy-monitoring or evidence-integration aspects of confidence. As such, they do not explain the integral affective and motivational aspects of confidence, nor appeal to appraisal mechanisms leveraged in affective science (Dukes et al., 2021; Moors et al., 2013; Scherer & Moors, 2019) that could account for how metacognitive affect is generated.

Appraisal theories in the affective-science literature broadly propose that affective valence reflects evaluations of goal-congruence and progress toward valued outcomes (Carver, 2015; Moors et al., 2021). This framework naturally extends to epistemic goals such as accuracy or understanding (Proust, 2014; Velasco & Loev, 2025). From this perspective, feeling confident corresponds to appraisal of congruence with (or progress towards) an epistemic goal, whereas doubt reflects goal incongruence (or insufficient progress). Finally, predictive-processing accounts elaborate this idea by embedding appraisal within a hierarchical generative model, where confidence is conceptualized as arising from prediction-error (non)-resolution and fluctuations in uncertainty (Friston et al., 2015; Hesp et al., 2021; Velasco & Loev, 2021, 2025).

Despite offering compelling conceptual links between confidence and affect, these accounts remain empirically and mechanistically underdeveloped. Although these accounts offer rich conceptual foundation, they lack mechanistic and empirical grounding needed to explain how metacognitive affect emerges during an ongoing decision processes. Existing empirical studies exploring affect in the context of metacognition typically rely on broad incidental affect manipulations such as mood inductions or affective priming manipulations (Culot et al., 2021; Culot & Gevers, 2021; Koellinger & Treffers, 2015; Massoni, 2014; Sidi et al., 2018). These type of manipulations do not necessarily support understanding how metacognitive affect emerges from task-related decision variables nor how it regulates cognition in an otherwise emotionally neutral context. Understanding how metacognitive affect emerges could therefore advance our understanding of the mechanisms of metacognition and move beyond viewing confidence solely as a representation of decision accuracy and more as a regulatory signal with affective antecedents and consequences.

Understanding these mechanisms has not only theoretical but also practical significance. In decision-making, confidence underpins belief updating (Rollwage et al., 2020) and information-seeking (Desender et al., 2018), with miscalibration linked to overconfidence in domains ranging from finance (Pikulina et al., 2017) to medicine (Berner & Graber, 2008). In education, students with well-calibrated confidence are more likely to choose effective learning strategies and perform better (David et al., 2024; Panadero et al., 2017). Distorted confidence also plays a role in mental health, where underconfidence maintains compulsive doubt in anxiety and OCD (Rouault et al., 2018), while overconfidence is implicated in delusional thinking in psychosis (Balzan, 2016) and inflated judgments in mania (Johnson & Jones, 2009). Finally, confidence is increasingly relevant for developing resilience to misinformation (Pennycook & Rand, 2021) and for effectively engaging with thinking-and decision-support systems such as large language models (Steyvers et al., 2025).

The overarching aim of this thesis is to investigate the mechanisms of metacognitive affect and move towards a perspective that integrates both cognitive and affective aspects of confidence. Through different studies, I explore the idea that confidence is a form of metacognitive affect emerging from predictive evaluations of one's accuracy. Focusing on the computational and neural mechanisms supporting this proposal, the thesis has three broader aims: (i) establish affect as metacognitive signal by testing whether it tracks decision accuracy and shares statistical structure with metacognitive confidence (Chapter 2); (ii) specify computational mechanisms of metacognitive affect by comparing goal-congruence with multilevel prediction-error accounts and instantiating a computational process model (Chapters 3 and 4); (iii) constrain the neural correlates of metacognitive signals by examining performance-monitoring ERPs (ERN/Pe) under difficulty and fluency manipulations (Chapter 5). These aims are operationalized into the following more specific research questions that structure the thesis:

Chapter 2: Do confidence and affect share similar statistical signatures reflecting decision accuracy? (Study I, Voodla et al., 2025) Here, we find that both confidence and metacognitive affect ratings reflect the subjective probability of being correct in a perceptual decision-making task, as they vary similarly when evidence for a decision or its expected difficulty is manipulated.

Chapter 3: Can we mechanistically explain metacognitive affect via (multilevel) prediction errors or appraisals of epistemic goal congruence? (Study II, Voodla et al., 2024; Study III Vrantisidis et al., 2025). We find that metacognitive affect tracks both epistemic goal congruence but also multilevel accuracy prediction errors in perceptual decisions, pointing towards a hybrid account.

Chapter 4: Can we leverage drift diffusion models (DDMs) to construct a process model of metacognitive affect generation? (Study II, Voodla et al., 2024). We implement and provide initial validation for AffectDDM, a DDM-based computational framework, where metacognitive affect is operationalized as the prediction error between expected and actual drift-rate for evidence accumulation in a perceptual decision.

Chapter 5: Do performance-monitoring ERPs explain variations in metacognitive signals such as confidence induced through manipulations of difficulty and fluency (Study IV, Voodla & Uusberg, 2021)? We find that Error-Related Negativity (ERN) covaries with confidence but does not explain difficulty/fluency effects on confidence. We also find that Error Positivity (Pe) does not reliably track confidence, implying a partial dissociation between implicit monitoring signals and explicit confidence reports.

2. CONFIDENCE AND AFFECT AS METACOGNITIVE SIGNALS

Our ability to evaluate the accuracy of our own decisions relies not only on external outcomes (i.e, whether a choice led to a good outcome such as reward) but also on our subjective internal signals, such as our sense of confidence or how positive or negative we feel (affective valence). These internal signals are often treated as separate constructs – confidence is typically defined as a cognitive estimate of decision accuracy, whereas affect in decision-making contexts is generally defined as a response to the (expected) outcomes of the decision (Emanuel & Eldar, 2023; Lerner et al., 2015; Phelps et al., 2014). However, recent theoretical accounts challenge this dichotomy, proposing that confidence is better defined constitutively as an affectively valenced feeling regulating the decision-making process (Velasco & Loev, 2025).

To clarify these proposals, it is useful to place confidence and epistemic affect within the broader context of metacognition research. Metacognition is usually parsed into monitoring and control functions (Flavell, 1979; Hennecke & Bürgler, 2023; Nelson, 1990). Monitoring refers to self-evaluation of current cognitive processes or outcomes, such as estimating if progress is made towards a goal or whether a decision is correct. Control refers to the adjustment of cognitive or behavioral strategies based on these assessments, for example gathering more evidence or sticking to the decision.

Both monitoring and control can be in turn decomposed into inputs and outputs (Table 1). Monitoring inputs refer to the signals or cues that the meta-cognitive system processes to evaluate ongoing decision processes (i.e. decision accuracy). Monitoring outputs are representations or experiences that are generated as consequence of this evaluation (i.e. feelings or cognitions). Control inputs are signals that carry information about the extent of regulation needed (i.e. confidence, affect) and control outputs are the behavioral or cognitive adjustments that follow. In the following, I will review monitoring inputs that have been empirically shown to drive confidence and metacognitive affect and control outputs that have been shown to be modulated by confidence or metacognitive affect.

Table 1. Inputs and outputs of confidence and (metacognitive) affect

Monitoring inputs	Construct	Control outputs
Evidence strength (Kiani & Shadlen, 2009)	Confidence	Information-seeking (Desender et al., 2018)
Response time (Ratcliff & Starns, 2009)		Effort allocation (Lee & Daunizeau, 2021)
Fluency (Oppenheimer, 2008)		Exploration vs exploitation (Boldt et al., 2019)
Expectations (Van Marcke et al., 2024)		Learning (Meyniel et al., 2015)
Errors (Yeung & Summerfield, 2012)		Prioritization (Aguilar-Lleyda et al., 2020)
Conflict (De Neys et al., 2011)		Changes of mind (van Den Berg et al., 2016)
Prediction errors (Emanuel & Eldar, 2023)	(Metacognitive) affect	Information-seeking (Yang & Kahlor, 2013)
Conflict (Dignath et al., 2020)		Effort allocation (Silvestrini & Gendolla, 2019)
Fluency (Reber et al., 1998)		Exploration vs exploitation (Van Dooren et al., 2021)
Uncertainty (Bar-Anan et al., 2009)		Learning (Eldar et al., 2016)
Goal congruence (Moors et al., 2013)		Prioritization (Carver, 2015)
Goal progress (Carver, 2015)		Control allocation (Dignath et al., 2020)

Confidence and metacognitive affect as monitoring outputs

Confidence is widely defined as a canonical monitoring output reflecting a subjective estimate of how likely a decision is to be correct (Fleming & Daw, 2017; Kepecs & Mainen, 2012). In this research direction, confidence has been shown to be modulated by multiple monitoring inputs, such as strength of evidence for an option (Kiani & Shadlen, 2009), response time (Ratcliff & Starns, 2009), prior expectations (Van Marcke et al., 2024), processing fluency (Oppenheimer, 2008), error likelihood (Yeung & Summerfield, 2012) and cognitive conflict (De Neys et al., 2011). Similarly, growing theoretical and empirical work treats certain affective states, such as curiosity, frustration, but also confidence as outputs of a metacognitive monitoring system (i.e. epistemic/metacognitive feelings; Arango-Muñoz, 2014; Fernández Velasco & Loev, 2025; Proust, 2014). These feelings

signal the ongoing status or outcome evaluations of cognitive processes such as decision-making or belief formation and recruit corrective control mechanisms if required. Across different contexts, researchers have described several metacognitive feelings such as the feeling of knowing (FOK), the feeling of “tip-of-the-tongue”, feeling of difficulty and feeling of rightness (FOR) (see Fernández Velasco & Loev, 2025 for an overview).

While these feelings vary in their specific content, they share a commonality as they are experienced as affectively valenced – i.e. they are experienced as positive or negative. This valence, rather than being an epiphenomenon, is proposed to constitute the felt component of metacognitive monitoring which drives regulatory consequences. Positive affect typically accompanies fluency and resolution, while negative valence arises from disfluency, uncertainty, or conflict (Velasco & Loev, 2025). Similarly to confidence, various monitoring inputs that drive affect in epistemic contexts have been characterized. For instance, prediction errors (Emanuel & Eldar, 2023), conflict (Dignath et al., 2020), fluency (Reber et al., 1998), uncertainty (Bar-Anan et al., 2009), goal congruence (Moors et al., 2013) and goal progress (Carver, 2015).

Confidence and affect as control inputs

Both confidence and affect are also well-established as control input signals. Confidence has been shown to act as a control input by modulating information-seeking (Desender et al., 2018), effort allocation (Lee & Daunizeau, 2021), exploration vs exploitation (Boldt et al., 2019), learning (Meyniel, 2020), task prioritization (Aguilar-Lleyda et al., 2020) and changes of mind (van Den Berg et al., 2016) among the key consequences. Similarly to confidence, affect has been shown to modulate information seeking (Yang & Kahlor, 2013), effort allocation (Silvestrini & Gendolla, 2019), exploration vs exploitation (Van Dooren et al., 2021), learning (Eldar et al., 2016), prioritization (Carver, 2015), control allocation (Dignath et al., 2020). Together, the evidence suggests that there is notable overlap between confidence and affect as inputs and outputs for a metacognitive monitoring and control system.

Taken together, this overlap suggests that metacognitive affect may play a role analogous to confidence. One way to test this is to ask whether metacognitive affect shows the same behavioral signatures that characterize confidence as a subjective representation of decision accuracy (Sanders et al., 2016), such as higher ratings for correct than error trials and increasing ratings with evidence strength. Study I was designed to test whether metacognitive affect tracks monitoring inputs that signal decision accuracy.

Study I

The goal of the Study I was to evaluate whether metacognitive affect, similarly to confidence, reflects participants' decision accuracy. To address this question, we used a perceptual decision-making paradigm, where participants had to perform binary decisions about the dominant direction of ambiguous visual stimuli. This task provides a controlled way to manipulate monitoring inputs such as evidence strength and expectancy, while minimizing interference from task-irrelevant emotional signals.

In study 1, two complementary experiments were conducted (Figure 1). In experiment 1, participants reported both confidence and affect sequentially on each trial, enabling direct comparison of how each report changes in response to manipulations of monitoring inputs. In experiment 2, confidence and affect responses were collected in separate task phases, enabling an assessment of whether these relationships persist even if they are not jointly elicited. Together, these experiments provide a robust test of whether confidence and metacognitive affect share accuracy monitoring signatures as outputs of a metacognitive monitoring system.

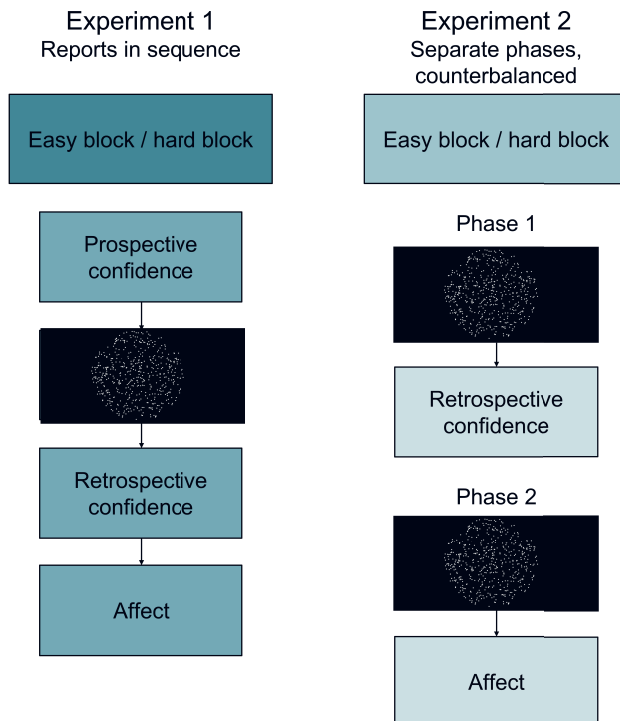


Figure 1. Experimental design for experiments 1 and 2. *In both experiments, we set up a random-dot motion task where participants had to evaluate the direction of the majority of dots in a dynamic stimulus and give a binary response (left or right). In Experiment I, 50 participants reported confidence and affect sequentially after a perceptual decision, in Experiment II, 39 participants reported either confidence or affect in*

separate phases of the experiment in counterbalanced order to rule out potential cross-contamination from joint reporting. In Experiment 1, all reports were given within the same trial in a fixed sequence: participants first provided a prospective-confidence estimate, then made the RDK direction decision, then reported retrospective confidence and affect. Confidence and affect were both reported on a 5-point scale (“Random guess” to “Certainly correct” for confidence and “Quite negative” to “Quite positive” for affect). Evidence strength had three coherence levels (.05 (hard), .20 (medium), .40 (easy)). Expectancy was manipulated by alternating easy and hard blocks of 15–20 trials, with block type inferred by the participant based on the recent trials. Participants were recruited online via Prolific and completed 360 trials each. In Experiment 2, reports of confidence and affect were separated into two counterbalanced phases so that each trial included only one report (Phase 1: decision with retrospective confidence; Phase 2: decision with affect). Scales were expanded to six points and, for confidence, explicitly spanned error to correct (i.e. Certainly wrong; Certainly correct as endpoints). There were two level of evidence strength (.05, .40) and expectancy was made explicit using mini-blocks of six trials with pre-trial and block cues. Participants were recruited via the KU Leuven Sona pool and completed 600 trials each.

The results of Study I indicated that both confidence and affect varied similarly as a response to monitoring input manipulations of decision accuracy, evidence strength, and expectancy (Figure 2). Confidence and affect displayed identical statistical signatures, increasing when decisions were more likely to be correct, had more evidence and when success was anticipated similarly to previously established statistical signatures of confidence (Sanders et al., 2016). Importantly, these effects were established in Experiment I and fully replicated in Experiment II, where confidence and affect were collected in separate task phases, ruling out the reporting order.

We found that confidence and affect ratings varied similarly in response to the manipulations targeting subjective probability that a decision is correct (Figure 2). First, both ratings increased with objectively correct responses, confirming their sensitivity to accuracy. Secondly, this increase was modulated by evidence strength for both confidence and affect, so that with more evidence, both ratings increased. Thirdly, we found an interaction between evidence strength and accuracy, so that confidence and affect ratings increased for correct trials with more evidence but were both not impacted by changes in evidence for error trials. Finally, both confidence and affect were sensitive to the expectancy manipulation, so that when participants expected a higher probability of success, their confidence and affect increased.

Study I demonstrates that metacognitive affect reflects the same monitoring inputs as confidence. Across two experiments, metacognitive affect reports showed the canonical statistical signatures of confidence, including sensitivity to decision accuracy, evidence strength, and expectancy. These effects persisted when confidence and affect were reported in separate task phases, ruling effects of sequential or joint reporting.

Together, these findings suggest that metacognitive affect reflects subjective estimates of decision accuracy rather than post-decisional emotional reactions. This aligns with accounts that treat epistemic feelings as outputs of metacognitive monitoring, and constrains theories of metacognitive affect in decision-making by showing that affective valence behaves as an accuracy-tracking epistemic signal.

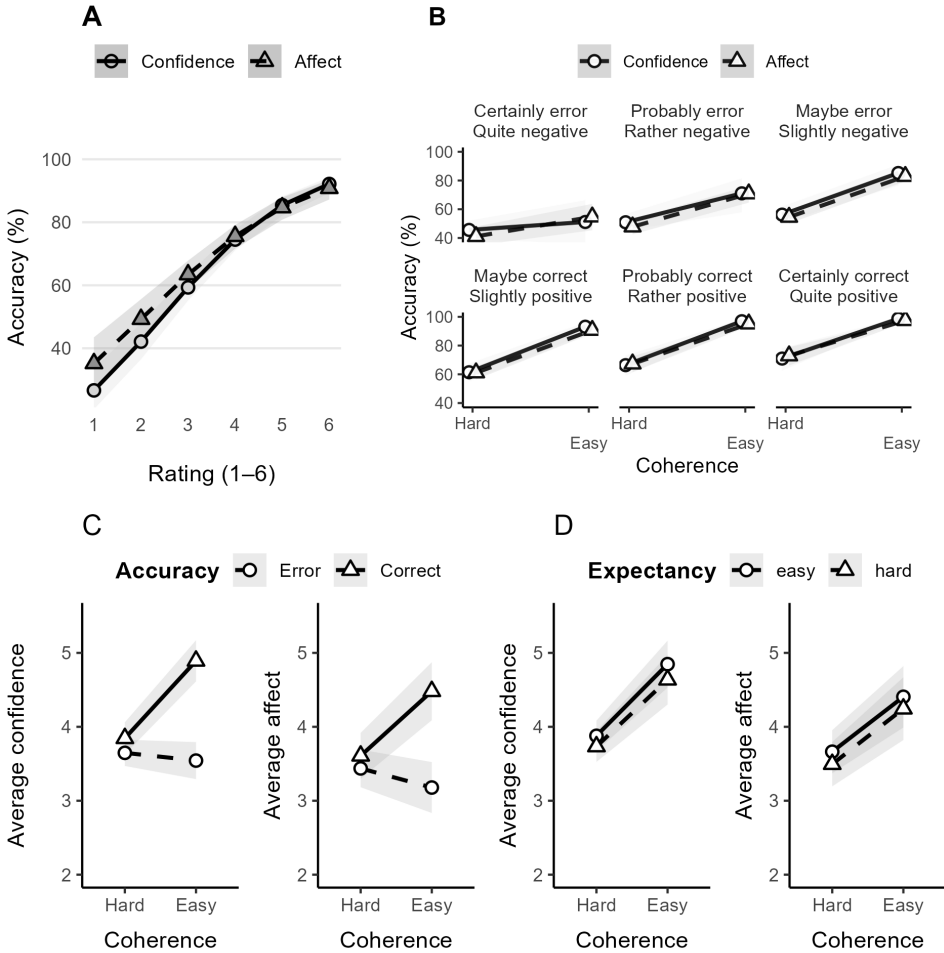


Figure 2. Statistical signatures of confidence from Voodla et al., 2025, Experiment 2. Panel A: confidence and affect increase linearly as function of accuracy. Panel B: at each level of confidence and affect, accuracy increases with evidence strength which indicates graded sensitivity to evidence for both confidence and affect reports. Panel C: Confidence and affect both increased with evidence strength for correct but not erroneous decisions. Panel D: confidence and affect are both similarly sensitive to expectations, so that expecting an easier trial leads to increased confidence/more positive affect. These results replicate the same patterns in Experiment 1 in the same paper. In all figures, averages reflect model-based marginal means, bands 95% confidence intervals.

3. MECHANISMS OF METACOGNITIVE AFFECT

The results presented in Chapter 2 supported the conclusion that metacognitive affect, like confidence, can function as a metacognitive monitoring signal tracking subjective accuracy. However, it is also important to understand the mechanisms through which metacognitive affect is generated. Understanding these mechanisms requires moving beyond functional descriptions towards an account of the particular computations that give rise to affective experiences. Broader theories of affect generation offer a natural starting point for this investigation as they specify how affective valence arises from computations of goal discrepancies, progress and/or expectation violations (Carver, 2015; Emanuel & Eldar, 2023; Hesp et al., 2021; Moors et al., 2021). When applied to metacognitive contexts, these theories make different predictions about which computations metacognitive affect should reflect. In particular, it is unclear whether metacognitive affect reflects discrepancy from an epistemic goal (i.e., goal discrepancy) or deviations from expectations about accuracy (i.e., prediction errors), or some combination of the two.

Two broad theoretical accounts can be leveraged to understand how metacognitive affect arises. According to appraisal theories and control-theoretic accounts of affect generation, affective valence reflects how effectively an individual is progressing towards a currently active goal (Carver & Scheier, 1990; Carver, 2015; Moors et al., 2017, 2021). According to this view, cues signaling forward progress generate positive affect, while cues indicating stalling or negative progress produce negative affect. Applied to epistemic contexts, this view predicts that affect should become more positive as decisions become easier or more accurate, regardless of prior expectations.

Prediction error accounts by contrast, propose that affective valence reflects discrepancies between expected and actual outcomes, i.e. rewards (Barrett, 2017; Emanuel & Eldar, 2023; Hesp et al., 2021; Joffily & Coricelli, 2013; Van de Cruys, 2017; Velasco & Loev, 2025). These accounts imply that affect is more positive when outcomes exceed expectations and more negative when they fall short. Applied to metacognitive affect, this view predicts that affect should be sensitive not only to actual progress toward being correct, but also to how that progress compares to what was expected. Consistent with this view, multiple studies outside epistemic contexts have found that trial-by-trial changes in affective valence (or momentary happiness) reflect reward prediction errors (Blain & Rutledge, 2020; Eldar et al., 2016; Emanuel & Eldar, 2023; Rutledge et al., 2014; Vanhasbroeck et al., 2021). Related mismatch effects have also been observed in perceptual and cognitive tasks, where violations of expectation such as surprise or conflict elicit affective responses (Chetverikov & Kristjánsson, 2016; Inzlicht et al., 2015).

Critically, these accounts diverge in their predictions about the role of expectations in shaping metacognitive affect. A pure goal-progress account predicts that affect should become more positive both when expected progress is high and when actual progress is high, as both signal effective movement toward the epistemic goal. Prediction-error accounts, by contrast, predict that lower

expectations should amplify positive affect, such that affect depends on the discrepancy between expected and actual progress rather than on progress alone. As a result, these accounts make distinguishable predictions about the main effects and interactions of expected and actual progress on affect. Discriminating between them requires experimental design that independently manipulates expected and actual progress while allowing their contributions to metacognitive affect to be dissociated.

A further complication is that expectations in decision-making tasks operate at multiple timescales (Cavalan et al., 2023; Van Marcke et al., 2024). Participants may hold local expectations about the probability of correct response on a specific trial, as well as more global expectations about performance within a task block or in the task as a whole. If prediction errors occur at both trial- and block levels, their combined influence may obscure the signature of prediction-error-based affect when analyses do not explicitly separate these levels.

The present chapter addresses these issues through two empirical studies. In Study II, we reused data from Experiment 2 of Study I and tested whether metacognitive affect reflects goal progress, expectations, or their interaction where trial difficulty (actual progress) and block-level expectancy (expected progress) were manipulated. Study III extended this design to explicitly disentangle trial-level and block-level prediction errors while controlling for temporal and contextual confounds. Together, these studies provide a principled test of whether metacognitive affect reflects goal progress, prediction errors, or a combination of both.

Study II

In Study II, we reanalyzed data from Experiment 2 of Study I to test whether metacognitive affect reflects goal progress, expectations, or their interaction. Participants performed a random-dot motion task in which actual progress was manipulated via trial-level motion coherence (easy vs. hard), while expected progress was manipulated via block-level difficulty (mostly easy vs. mostly hard blocks). After each decision, participants reported their affective valence on a 6-point scale.

Manipulation checks confirmed that task performance was strongly influenced by trial difficulty but not by block-level expectancy: participants were faster and more accurate on high- than low-coherence trials, with no main or interaction effects of expectancy on reaction time or accuracy. This indicates that our expectancy manipulation altered expectations about performance without directly affecting decision difficulty. The main analysis indicated that affect ratings varied additively as a function of both actual and expected progress (Figure 3). Affect was more positive on high- than low-coherence trials and more positive in easier than harder blocks, with no interaction between coherence and expectancy. Thus, affect was more positive when participants performed better and when they expected to perform well.

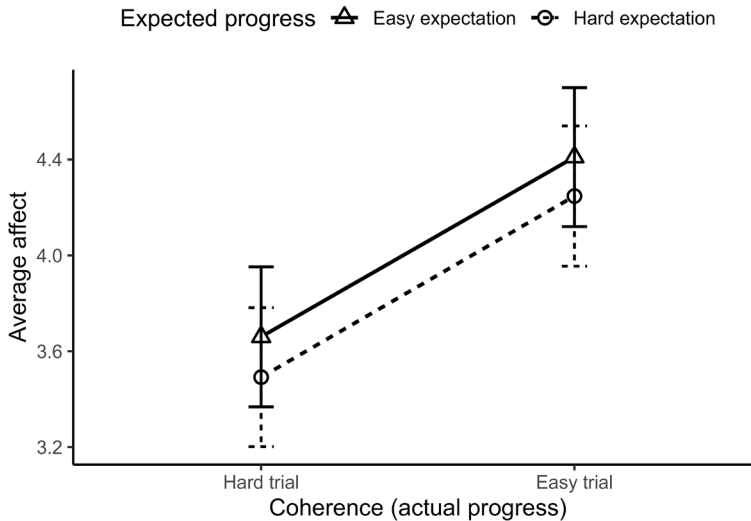


Figure 3. The effect of expected and actual progress on metacognitive affect. *Model-based estimated marginal means of metacognitive affect are shown as a function of trial-level motion coherence (actual progress; hard vs. easy trials) and block-level expectancy (expected progress; mostly hard vs. mostly easy blocks). Affect was rated on a 6-point scale, with higher values indicating more positive affect. Error bars indicate 95% confidence intervals.*

At face value, this additive pattern is consistent with a goal-progress account of metacognitive affect, according to which affect tracks progress toward the epistemic goal of being correct. However, this pattern is difficult to reconcile with a strict prediction-error account, which would instead predict more positive affect when expectations are low and outcomes are good. One possibility is that prediction errors operating at different timescales (i.e. block vs trial-level expectations) may have been conflated in this design, obscuring prediction-error effects at the trial level. This possibility motivated Study III, which was designed to more explicitly disentangle expectation-related signals across these two timescales.

Study III

Study III was designed to provide a stronger test of whether metacognitive affect reflects prediction errors by explicitly disentangling trial-level and block-level expectations. Whereas Study II manipulated expectations at the block level using short blocks of trials, Study III extended the design to separate expectation-related signals operating at different timescales. Participants performed a random-dot motion task similar to that used in Study II, but with several key modifications. Block length was increased substantially from 6 to 50 trials to allow block-level expectations to stabilize, block start labels were removed so that expectations had to be learned from experience with trials, and trial difficulty was varied continuously rather than categorically. Affect was measured on a

7-point valence scale with a neutral midpoint. These changes were designed to isolate prediction errors arising from moment-to-moment fluctuations in task difficulty from slower changes in contextual expectations.

Following the preregistered analysis plan, trial-by-trial affect ratings were analyzed using linear mixed-effects models that included measures of actual progress (trial difficulty) and expected progress (block difficulty), while controlling for block onset effects, feedback-related prediction errors, temporal trends, and affective carryover from previous trials. This approach ensured that trial-level prediction error effects were not conflated with block-level or task-related affective dynamics.

Manipulation checks confirmed that task performance was driven by trial-level coherence but not by block-level expectancy: participants were faster and more accurate on higher-coherence trials, with no main or interaction effects of expectancy on reaction time or accuracy. When replicating the analysis strategy used in Study II (affect ~ expected × actual progress), affect was more positive on easier trials, but showed neither a main effect of block expectancy nor an interaction between expectancy and coherence. Thus, with longer blocks, the positive expectancy effect observed in Study II did not replicate.

We next examined whether the effect of block-level expectancy varied systematically across trials within a block. We conducted an exploratory analysis including an interaction between block type and trial number in a block. This revealed a robust interaction, such that affect was more positive at the beginning of easy blocks, when participants initially expected trials to be harder but learned that they were actually easy which diminished as the block progressed. This increase in affect at block onset indicates a block-level prediction error that is resolved as participants update their expectations through experience. Given the short block length in Study II, this early block-level effect plausibly accounts for the additive expectancy effect observed there.

In the main pre-registered analysis, when explicitly controlling for block-onset effects, temporal trends, feedback-related prediction errors, and affective carryover, we found a pattern consistent with trial-level prediction errors. Affect remained more positive on easier trials and following correct responses, reflecting sensitivity to epistemic goal progress. Crucially, affect was also more positive in harder blocks than in easier blocks, indicating greater positive affect when expectations about performance were lower. This combination of effects provides evidence that metacognitive affect reflects both goal progress and discrepancies between expected and actual progress, consistent with a prediction-error-based account rather than an account based on goal progress alone.

Taken together, Studies II and III clarify how metacognitive affect is shaped by both epistemic goal progress and multilevel prediction errors. Study II showed that affect increased additively with trial difficulty and block-level expectancy, a pattern consistent with goal-progress accounts but seemingly inconsistent with prediction-error theories. However, because expectations were manipulated over short blocks, this design did not allow trial-level and block-level prediction errors to be separated.

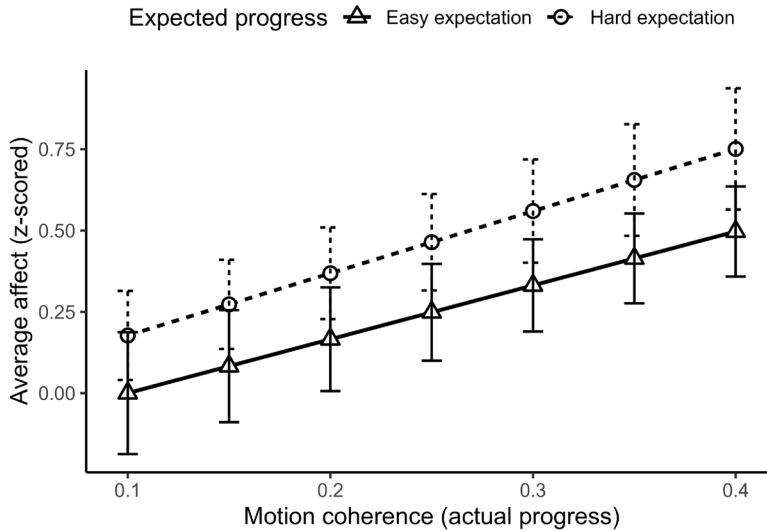


Figure 4. Effects of expected and actual progress on metacognitive affect in Study III. Model-based predictions of z-scored metacognitive affect are shown as a function of motion coherence (actual progress) for easy and hard blocks (expected progress). Points indicate estimated marginal means and error bars denote 95% confidence intervals.

Study III demonstrated that this apparent additivity of expected and actual progress reflects prediction errors operating at different timescales (Cavalan et al., 2023; Goekoop & de Kleijn, 2021; Hesp et al., 2021). When block-level expectations were allowed to stabilize and temporal confounds were controlled for, expectancy effects reversed in the direction predicted by prediction-error accounts: affect was more positive when performance exceeded lower expectations. At the same time, affect continued to track trial-level difficulty and accuracy, indicating sensitivity to epistemic goal congruence. These findings suggest that metacognitive affect does not reflect goal progress or prediction error in isolation pointing towards an integrative account (i.e. see Everaert et al., 2021 for a similar account). Instead, metacognitive affect appears to integrate information about how well one is doing and how that performance compares to what was expected, with prediction errors computed at multiple hierarchical levels contributing to the overall affective signal.

While the present results establish this integration at the behavioral level, they do not yet specify the computational mechanisms by which such signals are generated. Existing models of perceptual decision-making and confidence describe how evidence accumulation gives rise to choice accuracy and confidence estimates (Fleming & Daw, 2017; Shekhar & Rahnev, 2024), but they do not formalize how affective valence emerges in this process. Likewise, reinforcement-learning models formalize prediction errors over external rewards (i.e. (Blain & Rutledge, 2020; Eldar et al., 2016) but are not designed to capture generation of metacognitive affect.

These limitations motivate the need for a computational framework that can jointly account for evidence accumulation and metacognitive affect within a unified architecture. In particular, such a model should explain how prediction errors can be computed over internal decision variables, how these errors interact with absolute evidence strength, and how both contribute to affective experience. The next chapter addresses this challenge by introducing a computational model that formalizes metacognitive affect as an emergent signal derived from evidence accumulation, providing a mechanistic account of the behavioral patterns identified here.

4. AffectDDM – A COMPUTATIONAL FRAMEWORK OF METACOGNITIVE AFFECT

The results presented in the previous chapter indicated that both goal progress and prediction errors play a role in metacognitive affect generation. While these findings clarify drivers of metacognitive affect, they leave open a more fundamental question about how such affective signals are generated from the underlying decision processes. Existing computational models address parts of this problem in isolation: on the one hand, evidence accumulation models of perceptual metacognition (i.e. Fleming & Daw, 2017; Kepecs & Mainen, 2012; Shekhar & Rahnev, 2024) provide detailed accounts of how internal decision variables give rise to confidence judgments, but do not account for metacognitive affect generation. On the other hand, computational models of affect in decision-making have primarily focused on affect as a reinforcement signal (Emanuel & Eldar, 2023; Rutledge et al., 2014) typically driven by external outcomes rather than internal self-evaluations. As a result, we currently lack a computational framework that formalizes metacognitive affect as a dynamic, epistemic regulation signal that emerges from the decision process. Addressing this gap requires a modeling approach that can specify how affect emerges from internal decision variables as the decision unfolds, rather than being inferred post hoc from outcomes or behavior. This motivates the use of computational modeling as a tool for formalizing candidate mechanisms of metacognitive affect generation.

Computational modeling is a broad category of methods that enable quantitatively describing the processes underlying psychological phenomena (Farrell & Lewandowsky, 2010). Importantly, such models differ from standard data-analytic approaches (i.e., statistical or machine learning models), which aim to characterize relationships present in the observed data (Palminteri et al., 2017). By contrast, computational models aim to mathematically capture the generative mechanisms that produce these data in the first place (Piccinini & Craver, 2011). This distinction is crucial in domains where internal cognitive states responsible for observable behavior cannot be measured directly. By specifying and comparing candidate generative processes, computational models allow us to ask not only whether metacognitive affect correlates with decision accuracy, but how it might arise from interactions between internal variables of the system, for example, the desired level of evidence and time spent thinking on a decision. In this way, computational models instantiate candidate explanatory theories as simulations and force verbal theories to be more explicit about assumptions (Guest & Martin, 2021). As a consequence, computational models provide a bridge from verbal theories to precise testable predictions about the mechanisms responsible for the target phenomena.

In this chapter, I introduce a computational framework, AffectDDM, originally developed in Study II which extends evidence accumulation models from perceptual metacognition to metacognitive affect. The framework formalizes affect as an accuracy-monitoring signal grounded in evaluation of expected and actual progress in evidence accumulation during perceptual decisions. This model

provides a foundation for simulating the antecedents and consequences of affect as a metacognitive regulation signal, and thus serves as a methodological bridge toward more rigorous accounts of epistemic feelings.

The affectDDM framework

To understand the mechanisms of metacognitive affect generation, we implemented a computational framework that formalizes metacognitive affect within an evidence accumulation model. As an experimental paradigm, we focused on a perceptual decision-making task, where the dominant goal of the participant is to make accurate decisions. The central requirement for the model is that it should represent both actual progress during a decision and expected progress, allowing affect to be computed as their function. By specifying these computations explicitly, the model enables trial-by-trial predictions of affect and provides a rigorous way to test competing accounts of how metacognitive affect emerges during decision-making.

To develop the model, we needed a framework that parsimoniously characterizes the underlying decision-making process while linking affect generation to these underlying decision process. This has been successfully established in the perceptual metacognition literature, where the link between first-order (choice) and second-order (confidence) judgments has been well-described by extending evidence accumulation models, most prominently drift-diffusion models (DDMs) (Fleming & Daw, 2017; Kepecs & Mainen, 2012; Ratcliff et al., 2016). DDMs describe perceptual decisions as iterative sampling processes of noisy evidence which continue until a decision boundary is reached. In this framework, a perceptual decision process is governed by a small set of parameters that can be used to predict observable decision outcomes, such as choices, errors and reaction times. The model has also been extended to simultaneously account for confidence judgements as read-outs from accumulated evidence (Yeung & Summerfield, 2012). Crucially, DDMs provide a way to estimate latent subjective progress signal (drift-rate), which represents how efficiently evidence is accumulated. Following the findings in the previous chapter, this could be interpreted as the goal progress representation underlying metacognitive affect.

However, a computation of actual progress alone is insufficient. Following the findings from the previous chapter, metacognitive affect is also sensitive to predictions about progress (also see Blain & Rutledge, 2020; Emanuel & Eldar, 2023). This implies that the system should also represent how efficiently evidence is expected to accumulate. Extending this logic to perceptual decision-making implies that, alongside the actual drift rate indexing progress within a trial, the system should maintain an internal representation of expected drift-rate. Expected drift-rate reflects how quickly evidence should accumulate given task demands and recent experience. Metacognitive affect can then emerge from the comparison between observed and expected progress during the decision process.

To incorporate expectations, AffectDDM (Figure 5) extends the standard DDM with a parallel accumulator that represents expected progress estimated from prior block-level information (e.g. whether the current context is easy or hard).

This expectancy-driven process follows the same dynamics as the stimulus-driven accumulator but is governed by an expected drift rate parameter. Thus, for each trial, the model tracks both actual accumulated evidence and expected accumulated evidence. Both accumulators continue beyond the decision threshold for a fixed post-decisional time window, reflecting the fact that affect ratings are provided after the choice and may incorporate post-decisional information (Desender et al., 2021).

In the model, metacognitive affect is formalized as a function of estimated actual and expected progress at the time of giving the affect report. This signal is mapped onto the participant’s rating scale using individual bias and scaling parameters, allowing the model to generate trial-by-trial predictions of affect alongside standard behavioral measures such as choices and reaction times. Together, these features provide a unified computational framework for linking latent decision variables to metacognitive affect reports. In the next section, we provide an initial empirical validation of AffectDDM by testing whether the model reproduces key behavioral and affective regularities observed in the dataset of Study II. The full mathematical specification of the model is also presented in Study II.

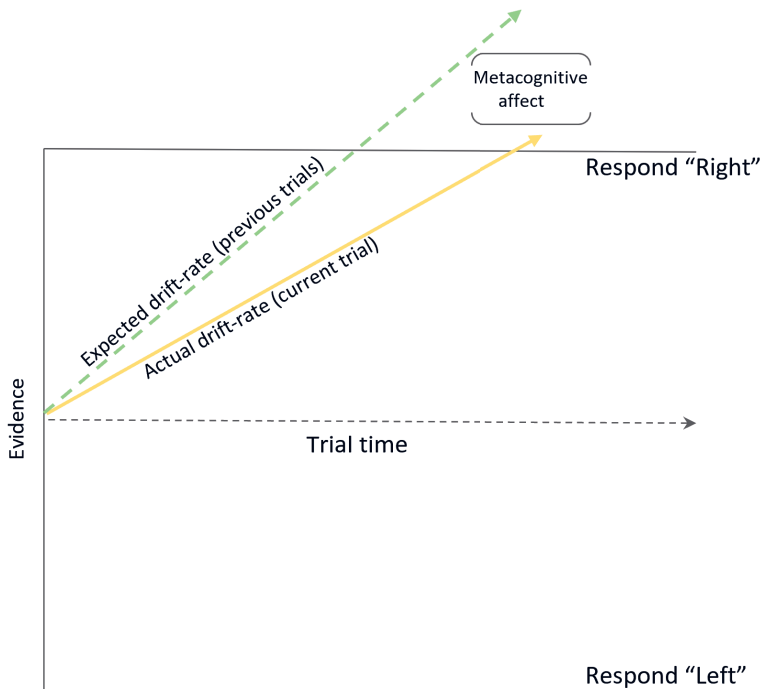


Figure 5. Metacognitive affect computation within the affectDDM framework. *The model includes two parallel evidence accumulation processes (coloured lines in the figure). The expectancy-driven accumulator (green dashed line) is governed by an expected drift rate estimated from previous trials, whereas the stimulus-driven accumulator (yellow line) is governed by the actual drift rate of the current trial. Metacognitive affect is computed from the discrepancy between observed and expected accumulated evidence at the time of affect rating.*

Initial validation of affectDDM

Here, I provide an initial behavioral validation of AffectDDM using the dataset introduced and analyzed from a behavioral perspective chapter 3. The goal is to assess whether the model reproduces the principal regularities in choices, reaction times, and affective responses under manipulations of actual progress and block-level expected progress. Using the same dataset as in the previous chapter, 39 participants performed a random-dot motion discrimination task with block-wise manipulations of actual and expected progress, reporting affect on a 6-point scale after each decision. The goal of the present analysis is not to re-establish the behavioral effects reported in the previous chapter, but to evaluate model adequacy. Accordingly, AffectDDM was fitted to the data and assessed on its ability to reproduce the main empirical regularities in choice accuracy, reaction times, and affective reports across experimental conditions. The model parameters were estimated per participant, using quantile optimization on a joint cost function combining reaction time quantiles for correct/error trials and the affect rating distributions. Behavioral analyses using mixed-effects regression models are reported in the previous chapter and serve here as the empirical benchmarks against which model predictions are evaluated.

We first showed that AffectDDM reproduces the main behavioral regularities observed in the data. Simulated reaction times from the fitted model closely matched the empirical patterns, reproducing the robust effect of stimulus coherence on reaction times while showing no effect of block-level expectancy and no coherence-expectancy interaction (Figure 6A). Model-predicted choice accuracy showed the same pattern, with higher accuracy on high-coherence trials and no influence of expectancy (Figure 6B).

We then analyzed whether AffectDDM reproduces the affective response patterns. Model-predicted affect ratings showed additive effects of both stimulus coherence and block-level expectancy, with higher affect ratings for easy trials and for easier blocks without interaction between the two. This pattern closely matched the empirical means observed in the original data (Figure 6C). Beyond mean-level effects, AffectDDM also reproduced the overall distribution of affect ratings across the six-point scale as well as the shape of reaction time distributions for correct and error.

Taken together, the results indicated that AffectDDM provides an adequate generative account of the observed behavioral and affective patterns in perceptual decision-making task. By embedding affect computation within the same evidence accumulation process that gives rise to choices and reaction times, the model captures how affective responses covary with internal decision variables rather than being tied solely to external outcomes.

In this dataset, affect was a positively associated with both actual and expected progress, and AffectDDM reproduces this structure by mapping accumulated evidence signals onto affective reports at the trial level. Future work could use AffectDDM to formally compare alternative implementations of affect computation, including models based on goal discrepancy or prediction error. Whereas

earlier analyses established that affect tracks both goal progress and prediction errors, the present model specifies how such signals can be generated within a unified computational architecture that has been widely used to model perceptual confidence. In this sense, AffectDDM complements the empirical findings of the previous chapter by providing a mechanistic account of how metacognitive affect can arise from internal monitoring of progress during decision-making.

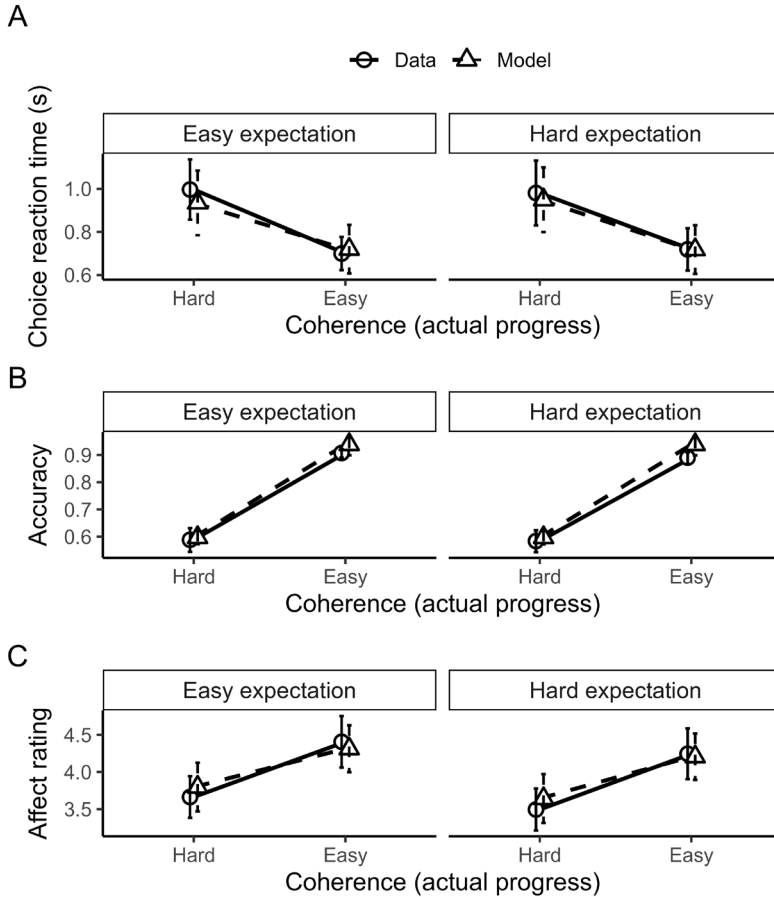


Figure 6. AffectDDM captures empirical effects of expected and actual progress. Panels depict choice reaction times (A), accuracies (B), and affect ratings (C) as a function of motion coherence (actual progress) and block expectancy (expected progress). Triangles and dashed lines show corresponding predictions from the AffectDDM. Points and error bars show marginal means estimated from empirical data or affectDDM predictions with 95% confidence intervals.

5. NEURAL CORRELATES OF METACOGNITIVE SIGNALS

The previous chapters conceptualized metacognitive affect and confidence as metacognitive signals: evaluative states that arise from internal monitoring of ongoing cognition and behavior and inform subsequent regulation. From this perspective, metacognitive signals integrate multiple sources of information about task demands, processing dynamics, and progress toward goals as shown in Chapter 2. Behavioral and computational approaches have been instrumental in clarifying what these signals track and how they may be generated at the algorithmic level. However, such approaches cannot directly inform us about the neural processes that support or constrain monitoring signals in the brain. Neural measures such as electroencephalography (EEG), in contrast, provide a window into the brain dynamics associated with performance monitoring processes that may contribute to metacognitive signals.

Research on neural correlates of performance-monitoring has identified a set of event-related potentials (ERPs) that are consistently elicited when participants evaluate or respond to their own performance (Ridderinkhof et al., 2004; Ullsperger et al., 2014; Yeung & Summerfield, 2012). Two of the most widely studied are the error-related negativity (ERN) (Falkenstein et al., 1991) and the error positivity (Pe) (Nieuwenhuis et al., 2001). The ERN is a sharp frontocentral negative deflection that emerges within ~100 ms of a response, and has been linked to activity in the anterior cingulate cortex. It is thought to reflect a fast mismatch signal indicating the deviation of outcomes from expectations, for example commission of errors. The Pe is a later positive deflection (~200–400 ms) with centroparietal distribution, often associated with the conscious awareness of errors with more recent integrative accounts proposing that Pe reflects post-decisional evidence accumulation (Desender et al., 2021).

It's important to note that these neural markers of performance monitoring need not correspond directly to subjective metacognitive experiences, but may also reflect implicit evaluation that only indirectly relate to metacognitive reports (Naccache et al., 2005; Shea et al., 2014). Performance-monitoring signals are typically fast, implicit, and automatically elicited, whereas metacognitive reports reflect constructed judgments that integrate multiple sources of information and may unfold over longer timescales. From this perspective, ERPs such as the ERN and Pe are better understood as indexing evaluative processes that provide inputs to metacognitive judgments, rather than constituting metacognitive judgements. As a result, neural monitoring activity may covary with metacognitive reports such as confidence under some conditions, while failing to account for how confidence is constructed or updated across different tasks.

Initial studies in perceptual decision-making tasks nevertheless suggested that ERN and Pe amplitudes covary with subjective confidence judgments, raising the possibility that they provide neural indices of computations supporting confidence confidence (Boldt & Yeung, 2015; Scheffers & Coles, 2000; Selimbeyoglu

et al., 2012). In these studies, larger Pe amplitudes have been linked to higher confidence, whereas larger ERN amplitudes have been associated with lower confidence, independently of objective accuracy. However, these findings have not always replicated (e.g. Rausch et al., 2020), and more recent work suggests that the relationship between performance-monitoring ERPs and confidence may be context-dependent rather than assumed.

Two further limitations of the existing literature are particularly relevant for the present work. First, it remains unclear whether the observed relationships between ERN, Pe, and confidence generalize beyond simple perceptual decisions to more complex reasoning tasks. Second, prior studies have largely relied on correlational designs and have not attempted to test whether experimental manipulations known to influence confidence such as decision fluency or task difficulty can be explained by corresponding changes in performance-monitoring ERPs. In addition, previous work has predominantly focused on error trials, treating the ERN and Pe as error-specific signals. If, however, these components reflect a more general performance-monitoring process related to confidence, they should covary with confidence irrespective of whether a response is correct or incorrect. This interpretation is supported by evidence for ERN-like activity on correct trials, often referred to as the correct-related negativity (CRN). Accordingly, in the present study we examine the relationship between confidence and both accuracy-specific (unpooled) and accuracy-collapsed (pooled) ERN and Pe components.

Study IV

The aim of Study IV was to test whether performance-monitoring ERPs account for changes in confidence during an arithmetic reasoning task, and whether they explain the effects of task difficulty and processing fluency on confidence. 48 participants performed a mental arithmetic task where they judged whether presented equations were correct or incorrect. Two within-subject factors were manipulated. Task difficulty was varied by presenting equations with either single-digit multiplicands (easy) or one double-digit multiplicand (hard). Response fluency was manipulated by briefly priming participants with either the correct or incorrect result prior to the presented equation. When the prime matched the correct solution, responses were expected to feel fluent; when it did not, responses were expected to feel disfluent. After each decision, participants reported their confidence on a visual analogue scale ranging from 0 to 100.

EEG was recorded using a 32-channel BioSemi ActiveTwo system and pre-processed using standard procedures including re-referencing, artifact correction, filtering, epoching relative to the response and baseline correction. The ERN was defined as mean frontocentral voltage (Fz, Cz) from -100 to 100 ms around the response, and the Pe as mean centroparietal voltage (Pz, Cz) from 175 to 325 ms post-response. However, if these components reflect a general performance-monitoring process that supports confidence, they should covary with confidence

irrespective of accuracy. Accordingly, in addition to standard accuracy-separated analyses, we also examined pooled ERN and Pe across correct and incorrect trials. Parallel unpooled analyses were also conducted.

Analyses proceeded in three steps – we first tested whether difficulty and fluency influenced behavioral performance and confidence. Second, we examined whether ERN and Pe were modulated by fluency and difficulty manipulations. Third, we used repeated-measures correlations and multilevel Bayesian mediation analyses to test whether ERN or Pe accounted for the effects of difficulty and fluency on confidence.

We found that Manipulations of task difficulty and response fluency produced the expected behavioral effects – participants were less accurate, slower, and less confident on hard compared to easy trials. Fluent trials were associated with faster responses and higher confidence, and their interaction indicated that the fluency effect on confidence was strongest for easy problems. These results confirmed that both manipulations successfully induced systematic variation in confidence.

At the neural level, task difficulty and fluency modulated ERN but not Pe, and these effects were expressed primarily during correct responding rather than error processing. When ERN was analyzed by pooling correct and incorrect trials, amplitudes were more negative for hard than for easy trials and for disfluent than for fluent trials (Figure 7A). Fluency and difficulty interaction indicated stronger fluency effects of ERN in easy conditions. In contrast, pooled Pe showed no effects of difficulty, fluency, nor their interaction. Separating correct and error trials we found no evidence of fluency or difficulty effects in error trials but robust effects of both in correct trials (Figure 7B). Pe showed no effects in error trials and a small effect of difficulty in correct trials.

Finally, we tested whether performance-monitoring ERPs mediated the effects of task difficulty and response fluency on decision confidence. Because Pe was neither sensitive to the experimental manipulations nor correlated with confidence, mediation analyses were conducted only for ERN. Using multilevel Bayesian mediation models, we examined whether trial-by-trial variations in pooled ERN or unpooled ERN accounted for the effects of difficulty and fluency on confidence. These analyses provided no evidence for mediation. Although difficulty and fluency influenced both confidence and ERN, changes in pooled/unpooled ERN did not explain how these manipulations translated into changes in confidence. Given the inherent noise of single-trial ERP measures, these null mediation effects should be interpreted cautiously, as the study may have been underpowered to detect small indirect effects.

We therefore found that across these analyses, task difficulty and response fluency reliably shaped confidence in arithmetic reasoning, whereas their neural signatures were only partially captured by classical performance-monitoring ERPs. ERN tracked both manipulations and covaried with confidence, but this relationship was expressed primarily during correct responding and did not mediate the behavioral effects of difficulty or fluency on confidence. Pe, in contrast, was not sensitive to the manipulations nor correlated with confidence.

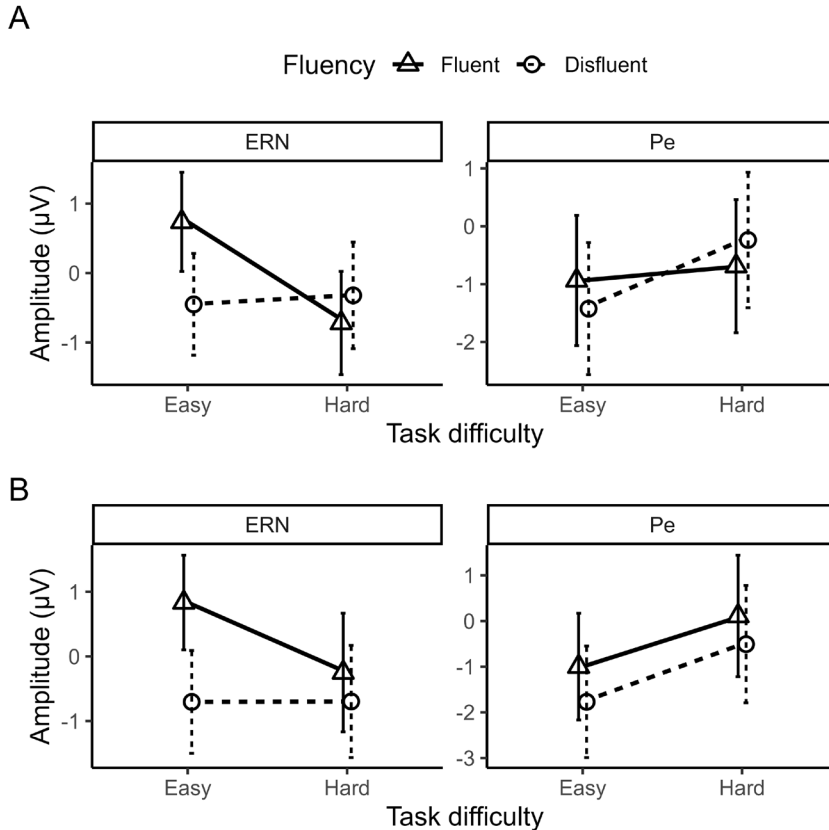


Figure 7. Effects of fluency and difficulty manipulations on ERN and PE in pooled and only correct trials. ERN amplitudes were more negative on hard compared to easy trials and on disfluent compared to fluent trials (A), with these effects primarily driven by correct responses (B). Pe amplitudes were not sensitive to either manipulation in pooled condition (A) but were sensitive to difficulty for only correct trials. Note that ERN amplitudes are plotted on the original (non-inverted) voltage scale. Points show estimated marginal means with 95% confidence intervals.

These results suggest that performance-monitoring activity indexed by ERN could reflect an implicit monitoring process than a neural implementation of confidence itself. These findings demonstrate that in future research neural correlates of metacognitive signals should be assumed to be context-dependent and differ substantially between simple perceptual decisions and more complex forms of reasoning.

6. SUMMARY AND CONCLUSIONS

The central aim of this thesis was to extend our understanding of metacognitive confidence with an affective perspective. While confidence has traditionally been treated as a judgment about the probability of being correct, everyday experience suggests that confidence also has a felt affective quality. The central question motivating this work was whether this affective dimension is incidental to confidence, or whether it reflects core mechanisms through which metacognitive monitoring operates. Across four empirical studies combining behavioral experiments, computational modeling, and neural measures, the thesis advances an affective perspective on metacognition. Rather than treating confidence and affect as separate responses to decision outcomes, the findings converge on the idea that both function as monitoring signals generated during ongoing cognition. In Study 1, confidence and affect were shown to track the same epistemic variables, including accuracy, evidence strength, and expectations. This parallelism suggests that confidence belongs to a broader class of metacognitive feelings that reflect progress toward epistemic goals.

At the mechanistic level, Studies II and III provide evidence that metacognitive affect reflects both epistemic goal progress and multilevel epistemic prediction errors. Building on this insight, the AffectDDM framework formalized affect as a dynamic signal arising from expected and actual progress within an evidence accumulation framework. While this model provides a concrete process-level account, it remains deliberately minimalist and does not attempt to model other components of affect generation or the downstream regulatory consequences of epistemic feelings, which is a fruitful future direction.

In Study IV, analysis of performance-monitoring ERPs showed limited and context-dependent relationships with confidence diverging from classical error-monitoring accounts. These findings suggest that error-monitoring signals such as the ERN may reflect implicit evaluative processes that do not straightforwardly map onto explicit confidence reports and caution against treating any single neural marker as a direct proxy for subjective confidence.

The findings of this thesis support understanding confidence as an affective epistemic signal rather than a purely cognitive estimate of accuracy. Approaching confidence from this angle implies that it should be further studied by drawing on existing theories of affect/emotion generation and self-regulation, and by asking different kinds of mechanistic questions than those that dominate the current literature focused on accuracy calibration. If confidence is affective, then it can be analyzed within appraisal-based accounts of affect generation (e.g. Moors et al., 2013; see also Erdemli et al., 2025 for a similar approach with curiosity). As demonstrated in the current thesis, metacognitive affect arises from evaluations of progress toward epistemic goals such as accuracy or understanding. Future work can therefore examine how different epistemic appraisals give rise to distinct metacognitive processes (e.g. re-evaluating the decision (Elosegi et al., 2024), gathering more evidence (Desender et al., 2018), adjusting epistemic

goals/thresholds (Balsdon et al., 2020; Hausmann & Läge, 2008), etc), and how these processes in turn support or hinder accurate self-evaluation.

An affective view of confidence also highlights the relevance of componential models of emotion (Scherer & Moors, 2019). If epistemic feelings are affective states, they should not be studied in isolation, but in relation to other emotion components, including motivational tendencies, physiological responses, expressive behavior, and subjective experience. Mapping how metacognitive affect covaries with these components would provide a more complete account of the mechanisms through which epistemic feelings emerge and influence cognition, including their embodied aspects (Barrett, 2017; Seth & Friston, 2016). Finally, conceptualizing confidence as affective naturally situates it within a self-regulatory framework (Gross, 2015). Affective states are not only experienced but also explicitly regulated, and feelings of high confidence or doubt can be desired or undesired depending on context (Efklides, 2011). From this perspective, confidence is also a target of emotion regulation and not only a readout of evidence. Future research can therefore examine how people attempt to up- or down-regulate metacognitive affect using strategies analogous to those studied in the emotion regulation literature, and illuminate how these regulatory attempts shape learning, decision-making, and longer-term belief formation.

In conclusion, the current thesis substantiates an affective perspective on metacognitive confidence. By examining metacognitive affect and confidence at behavioral, computational and neural levels, I demonstrate that confidence reflects context-sensitive affective processes tied to epistemic goals, rather than solely judgments of accuracy.

ACKNOWLEDGEMENTS

I am deeply grateful to Andero and Kobe for their steady support over the years and for inspiring me to do science with rigor and care but also with ease and humor.

I thank my colleagues at the University of Tartu and the TARU lab, especially Sille-Liis, Heidi, Kaijo, Siim, Helen, Martin, and Richard. I am equally thankful to members of DesenderLab and KU Leuven Brain & Cognition – especially Robin, Pierre, Helene, Gaia, and Massimo. I am also grateful for my experience in SPL Lab at Stanford University – thank you James, Pille-Riin, Kaitlyn, Victoria, Kathrin, David, Maya, Matt, Jinxiao, Ashish, and Kate, for inspiring discussions and a generous research environment. I am also grateful to Thalia for pushing the ideas presented in the thesis further.

I am grateful to Kertu and Marina for essential administrative support, and to the Kristjan Jaak scholarship for financial backing.

I thank my progress committee, Jaan and Agnes, for their thoughtful feedback and guidance.

I am grateful to my thinking partner Kaspar for never blocking the way of the inquiry.

Finally, I'm thankful to my two families (Johann, Mika, Piret, Tiit, Terje, Taaniel, Johanna, Alden, Frida) and above all to Kata for always being there for me.

REFERENCES

- Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences*, 21(8), 607–617.
- Aguilar-Lleyda, D., Lemarchand, M., & De Gardelle, V. (2020). Confidence as a Priority Signal. *Psychological Science*, 31(9), 1084–1096.
<https://doi.org/10.1177/0956797620925039>
- Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology*, 27(2), 193–211. <https://doi.org/10.1080/09515089.2012.732002>
- Balsdon, T., Wyart, V., & Mamassian, P. (2020). Confidence controls perceptual evidence accumulation. *Nat. Commun.* 11, 1753.
- Balzan, R. P. (2016). Overconfidence in psychosis: The foundation of delusional conviction? *Cogent Psychology*, 3(1), 1135855.
<https://doi.org/10.1080/23311908.2015.1135855>
- Bar-Anan, Y., Wilson, T. D., & Gilbert, D. T. (2009). The feeling of uncertainty intensifies affective reactions. *Emotion*, 9(1), 123.
- Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12(1), 1–23.
- Berner, E. S., & Graber, M. L. (2008). Overconfidence as a cause of diagnostic error in medicine. *The American Journal of Medicine*, 121(5), S2–S23.
- Blain, B., & Rutledge, R. B. (2020). Momentary subjective well-being depends on learning and not reward. *Elife*, 9, e57977.
- Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness*, 2019(1), niz004.
- Boldt, A., & Yeung, N. (2015). Shared neural markers of decision confidence and error detection. *Journal of Neuroscience*, 35(8), 3478–3484.
- Carver, C. S. (2015). Control Processes, Priority Management, and Affective Dynamics. *Emotion Review*, 7(4), 301–307. <https://doi.org/10.1177/1754073915590616>
- Cavalan, Q., Vergnaud, J.-C., & De Gardelle, V. (2023). From local to global estimations of confidence in perceptual decisions. *Journal of Experimental Psychology: General*, 152(9), 2544.
- Chetverikov, A., & Kristjánsson, Á. (2016). On the joys of perceiving: Affect as feedback for perceptual predictions. *Acta Psychologica*, 169, 1–10.
- Culot, C., Fantini-Hauwel, C., & Gevers, W. (2021). The influence of sad mood induction on task performance and metacognition. *Quarterly Journal of Experimental Psychology*, 74(9), 1605–1614. <https://doi.org/10.1177/17470218211004205>
- Culot, C., & Gevers, W. (2021). Happy is easy: The influence of affective states on cognitive control and metacognitive reports. *Cognition and Emotion*, 35(6), 1195–1202. <https://doi.org/10.1080/02699931.2021.1932427>
- David, L., Biwer, F., Baars, M., Wijnia, L., Paas, F., & De Bruin, A. (2024). The Relation Between Perceived Mental Effort, Monitoring Judgments, and Learning Outcomes: A Meta-Analysis. *Educational Psychology Review*, 36(3), 66.
<https://doi.org/10.1007/s10648-024-09903-z>
- De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PloS One*, 6(1), e15954.
- 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>

- Desender, K., Donner, T. H., & Verguts, T. (2021). Dynamic expressions of confidence within an evidence accumulation framework. *Cognition*, *207*, 104522.
- Dignath, D., Eder, A. B., Steinhäuser, M., & Kiesel, A. (2020). Conflict monitoring and the affective-signaling hypothesis—An integrative review. *Psychonomic Bulletin & Review*, *27*(2), 193–216. <https://doi.org/10.3758/s13423-019-01668-9>
- Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., Broomhall, S., Brosch, T., Campos, J. J., & Clay, Z. (2021). The rise of affectivism. *Nature Human Behaviour*, *5*(7), 816–820.
- Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, *1*(1), 3–14.
- Efklides, A. (2011). Interactions of Metacognition With Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educational Psychologist*, *46*(1), 6–25. <https://doi.org/10.1080/00461520.2011.538645>
- Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as representation of momentum. *Trends in Cognitive Sciences*, *20*(1), 15–24.
- Elosegi, P., Rahnev, D., & Soto, D. (2024). Think twice: Re-assessing confidence improves visual metacognition. *Attention, Perception, & Psychophysics*, *86*(2), 373–380. <https://doi.org/10.3758/s13414-023-02823-0>
- Emanuel, A., & Eldar, E. (2023). Emotions as computations. *Neuroscience & Biobehavioral Reviews*, *144*, 104977.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455.
- Farrell, S., & Lewandowsky, S. (2010). Computational Models as Aids to Better Reasoning in Psychology. *Current Directions in Psychological Science*, *19*(5), 329–335. <https://doi.org/10.1177/0963721410386677>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive—developmental inquiry. *American Psychologist*, *34*(10), 906.
- Fleming, S. M. (2024). Metacognition and Confidence: A Review and Synthesis. *Annual Review of Psychology*, *75*(1), 241–268. <https://doi.org/10.1146/annurev-psych-022423-032425>
- Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review*, *124*(1), 91.
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1338–1349. <https://doi.org/10.1098/rstb.2011.0417>
- Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214. <https://doi.org/10.1080/17588928.2015.1020053>
- Goekoop, R., & de Kleijn, R. (2021). How higher goals are constructed and collapse under stress: A hierarchical Bayesian control systems perspective. *Neuroscience & Biobehavioral Reviews*, *123*, 257–285.
- Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. *Neuroscience & Biobehavioral Reviews*, *55*, 88–97.
- Gross, J. J. (2015). Emotion Regulation: Current Status and Future Prospects. *Psychological Inquiry*, *26*(1), 1–26. <https://doi.org/10.1080/1047840X.2014.940781>

- Guest, O., & Martin, A. E. (2021). How Computational Modeling Can Force Theory Building in Psychological Science. *Perspectives on Psychological Science*, 16(4), 789–802. <https://doi.org/10.1177/1745691620970585>
- Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision Making*, 3(3), 229–243.
- Hennecke, M., & Bürgler, S. (2023). Metacognition and self-control: An integrative framework. *Psychological Review*, 130(5), 1262.
- Hesp, C., Smith, R., Parr, T., Allen, M., Friston, K. J., & Ramstead, M. J. (2021). Deeply felt affect: The emergence of valence in deep active inference. *Neural Computation*, 33(2), 398–446.
- Hoven, M., Lebreton, M., Engelmann, J. B., Denys, D., Luigjes, J., & van Holst, R. J. (2019). Abnormalities of confidence in psychiatry: An overview and future perspectives. *Translational Psychiatry*, 9(1), 268.
- Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, 19(3), 126–132.
- Joffily, M., & Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Computational Biology*, 9(6), e1003094.
- Johnson, S. L., & Jones, S. (2009). Cognitive correlates of mania risk: Are responses to success, positive moods, and manic symptoms distinct or overlapping? *Journal of Clinical Psychology*, 65(9), 891–905. <https://doi.org/10.1002/jclp.20585>
- Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1322–1337. <https://doi.org/10.1098/rstb.2012.0037>
- Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210), 227–231.
- Kiani, R., & Shadlen, M. N. (2009). Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex. *Science*, 324(5928), 759–764. <https://doi.org/10.1126/science.1169405>
- Koellinger, P., & Treffers, T. (2015). Joy leads to overconfidence, and a simple countermeasure. *PloS One*, 10(12), e0143263.
- Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, 9(2), 149–171.
- Lee, D. G., & Daunizeau, J. (2021). Trading mental effort for confidence in the metacognitive control of value-based decision-making. *Elife*, 10, e63282.
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, 66(1), 799–823. <https://doi.org/10.1146/annurev-psych-010213-115043>
- Loev, S. (2022). Epistemic Feelings are Affective Experiences. *Emotion Review*, 14(3), 206–216. <https://doi.org/10.1177/17540739221104464>
- Massoni, S. (2014). Emotion as a boost to metacognition: How worry enhances the quality of confidence. *Consciousness and Cognition*, 29, 189–198.
- Meyniel, F. (2020). Brain dynamics for confidence-weighted learning. *PLOS Computational Biology*, 16(6), e1007935.
- Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, 88(1), 78–92.

- Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal Theories of Emotion: State of the Art and Future Development. *Emotion Review*, 5(2), 119–124. <https://doi.org/10.1177/1754073912468165>
- Moors, A., Van de Cruys, S., & Pourtois, G. (2021). Comparison of the determinants for positive and negative affect proposed by appraisal theories, goal-directed theories, and predictive processing theories. *Current Opinion in Behavioral Sciences*, 39, 147–152.
- Naccache, L., Dehaene, S., Cohen, L., Habert, M.-O., Guichart-Gomez, E., Galanaud, D., & Willer, J.-C. (2005). Effortless control: Executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia*, 43(9), 1318–1328.
- Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. In *Psychology of learning and motivation* (Vol. 26, pp. 125–173). Elsevier. <https://www.sciencedirect.com/science/article/pii/S0079742108600535>
- Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, 38(5), 752–760.
- Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, 12(6), 237–241.
- Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, 21(6), 425–433.
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74–98.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402.
- Phelps, E. A., LeDert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience*, 37(1), 263–287. <https://doi.org/10.1146/annurev-neuro-071013-014119>
- Piccinini, G., & Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, 183(3), 283–311.
- Pikulina, E., Renneboog, L., & Tobler, P. N. (2017). Overconfidence and investment: An experimental approach. *Journal of Corporate Finance*, 43, 175–192.
- Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, 19(3), 366–374.
- Proust, J. (2014). The representational structure of feelings. In *Open mind*. Open MIND. Frankfurt am Main: MIND Group. https://open-mind.net/papers/the-representational-structure-of-feelings/at_download/paperPDF
- 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.
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, 116(1), 59.
- Rausch, M., Zehetleitner, M., Steinhauser, M., & Maier, M. E. (2020). Cognitive modelling reveals distinct electrophysiological markers of decision confidence and error monitoring. *NeuroImage*, 218, 116963.
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of Perceptual Fluency on Affective Judgments. *Psychological Science*, 9(1), 45–48. <https://doi.org/10.1111/1467-9280.00008>
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The Role of the Medial Frontal Cortex in Cognitive Control. *Science*, 306(5695), 443–447. <https://doi.org/10.1126/science.1100301>

- Rollwage, M., Loosen, A., Hauser, T. U., Moran, R., Dolan, R. J., & Fleming, S. M. (2020). Confidence drives a neural confirmation bias. *Nature Communications*, *11*(1), 2634.
- Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biological Psychiatry*, *84*(6), 443–451.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145.
- Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, *111*(33), 12252–12257. <https://doi.org/10.1073/pnas.1407535111>
- Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a statistical computation in the human sense of confidence. *Neuron*, *90*(3), 499–506.
- Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 141.
- Scherer, K. R., & Moors, A. (2019). The Emotion Process: Event Appraisal and Component Differentiation. *Annual Review of Psychology*, *70*(1), 719–745. <https://doi.org/10.1146/annurev-psych-122216-011854>
- Selimbeyoglu, A., Keskin-Ergen, Y., & Demiralp, T. (2012). What if you are not sure? Electroencephalographic correlates of subjective confidence level about a decision. *Clinical Neurophysiology*, *123*(6), 1158–1167.
- Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160007. <https://doi.org/10.1098/rstb.2016.0007>
- Sharot, T., Rollwage, M., Sunstein, C. R., & Fleming, S. M. (2023). Why and When Beliefs Change. *Perspectives on Psychological Science*, *18*(1), 142–151. <https://doi.org/10.1177/17456916221082967>
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Sciences*, *18*(4), 186–193.
- Shekhar, M., & Rahnev, D. (2024). How do humans give confidence? A comprehensive comparison of process models of perceptual metacognition. *Journal of Experimental Psychology: General*, *153*(3), 656.
- Sidi, Y., Ackerman, R., & Erez, A. (2018). Feeling happy and (over)confident: The role of positive affect in metacognitive processes. *Cognition and Emotion*, *32*(4), 876–884. <https://doi.org/10.1080/02699931.2017.1346501>
- Silvestrini, N., & Gendolla, G. H. (2019). Affect and cognitive control: Insights from research on effort mobilization. *International Journal of Psychophysiology*, *143*, 116–125.
- Steyvers, M., Tejada, H., Kumar, A., Belem, C., Karny, S., Hu, X., Mayer, L. W., & Smyth, P. (2025). What large language models know and what people think they know. *Nature Machine Intelligence*, *7*(2), 221–231.
- Thompson, V. A., Turner, J. A. P., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140.
- Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of Performance Monitoring and Adaptive Behavior. *Physiological Reviews*, *94*(1), 35–79. <https://doi.org/10.1152/physrev.00041.2012>

- Van de Cruys, S. (2017). *Affective value in the predictive mind*.
<https://lirias.kuleuven.be/retrieve/443336>
- van Den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. *Elife*, 5, e12192.
- Van Dooren, R., De Kleijn, R., Hommel, B., & Sjoerds, Z. (2021). The exploration-exploitation trade-off in a foraging task is affected by mood-related arousal and valence. *Cognitive, Affective, & Behavioral Neuroscience*, 21(3), 549–560.
<https://doi.org/10.3758/s13415-021-00917-6>
- Van Marcke, H., Denmat, P. L., Verguts, T., & Desender, K. (2024). Manipulating Prior Beliefs Causally Induces Under- and Overconfidence. *Psychological Science*, 35(4), 358–375. <https://doi.org/10.1177/09567976241231572>
- Vanhasbroeck, N., Devos, L., Pessers, S., Kuppens, P., Vanpaemel, W., Moors, A., & Tuerlinckx, F. (2021). Testing a computational model of subjective well-being: A preregistered replication of Rutledge et al. (2014). *Cognition and Emotion*, 35(4), 822–835. <https://doi.org/10.1080/02699931.2021.1891863>
- Velasco, P. F., & Loev, S. (2021). Affective experience in the predictive mind: A review and new integrative account. *Synthese*, 198(11), 10847–10882.
<https://doi.org/10.1007/s11229-020-02755-4>
- Velasco, P. F., & Loev, S. (2025). Metacognitive Feelings: A Predictive-Processing Perspective. *Perspectives on Psychological Science*, 20(4), 691–713.
<https://doi.org/10.1177/17456916231221976>
- Voodla, A., & Uusberg, A. (2021). Do performance-monitoring related cortical potentials mediate fluency and difficulty effects on decision confidence? *Neuropsychologia*, 155, 107822.
- Voodla, A., Uusberg, A., & Desender, K. (2024). Affective valence does not reflect progress prediction errors in perceptual decisions. *Cognitive, Affective, & Behavioral Neuroscience*, 24(1), 60–71. <https://doi.org/10.3758/s13415-023-01147-8>
- Voodla, A., Uusberg, A., & Desender, K. (2025). Metacognitive confidence and affect – two sides of the same coin? *Cognition and Emotion*, 1–18.
<https://doi.org/10.1080/02699931.2025.2451795>
- Vrantsidis, T., Voodla, A., & Sabbagh, K. (2025). *Reward Prediction Errors, Not Expectations Or Outcomes, Drive Emotional Valence*. Dataset]. [osf. io/v5ubg](https://osf.io/v5ubg).
<https://osf.io/upjw4/download>
- Yang, Z. J., & Kahlor, L. (2013). What, Me Worry? The Role of Affect in Information Seeking and Avoidance. *Science Communication*, 35(2), 189–212.
<https://doi.org/10.1177/1075547012441873>
- Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1310–1321. <https://doi.org/10.1098/rstb.2011.0416>

KOKKUVÕTE

Kas see tundub õige? Afektiivne vaade metakognitiivsele otsustuskindlusele

Käesoleva doktoritöö eesmärk oli laiendada arusaama metakognitiivsest kindlusest otsutamise kontekstis, käsitledes seda afektiivteaduse vaatenurgast. Traditsiooniliselt on otsusekindlust mõistetud hinnanguna otsuse õigsuse tõenäosusele, kuid igapäevane kogemus viitab, et kindlusel on ka afektiivne mõõde. Töö keskne küsimus oli, kas see afektiivne aspekt on otsusekindluse juhuslik kõrvalnähtus või peegeldab see metakognitiivse enesemonitoorimise kesket mehhanismi.

Doktoritöö koosneb neljast empiirilisest uuringust, milles kombineeritakse käitumuslikke eksperimente, komputatsioonilist modelleerimist ja neurofüsioloogilisi mõõtmisi. Tulemused näitavad, et otsusekindlus ja afekt toimivad sarnaste metakognitiivsete signaalidena, mis tekivad otsustusprotsessi käigus. Esimeses uuringus leiti, et otsustuskindlus ja afekt on tundlikud samadele mõjutustele otsustusprotsessis, sealhulgas otsuse õigsusele, tõendite tugevusele ja ootustele, viidates sellele, et otsustuskindlus kuulub laiemasse metakognitiivsete tunnete klassi, mis peegeldavad progressi epistemiliste eesmärkide suunas.

Mehhanistlikul tasandil näitavad teise ja kolmanda uuringu tulemused, et metakognitiivne afekt peegeldab nii edenemist episteemiliste eesmärkide suhtes kui ka mitmetasandilisi epistemilisi ennustusvigu. Nende mehhanismide täpsemaks kirjeldamiseks töötati välja afektiivne otsuse triivmudel (AffectDDM), mis formaliseerib afekti kujunemise dünaamilise signaalina oodatud ja tegeliku progressi võrdlusele. Neljandas uuringus leiti, et soorituse monitoorimisega seotud EEG sündmuspotentsiaalide seosed otsustuskindlusega on kontekstist sõltuvad, mis viitab lahknevusele metakognitsiooni neuraalsete signaalide ja teadlikult raporteeritud otsustuskindluse vahel.

Kokkuvõttes toetavad käesoleva doktoritöö tulemused arusaama, et metakognitiivne otsustuskindlus on oma olemuselt afektiivne nähtus. Metakognitiivse afekti ja otsustuskindluse uurimine käitumuslikul, arvutuslikul ja neuraalsel tasandil näitab, et otsustuskindlus peegeldab kontekstitundlikku afektiivset protsessi, mis on seotud laiemalt episteemiliste eesmärkide ja ootustega ning mitte üksnes otsuse õigsuse kognitiivsete hinnangutega.

SAMENVATTING

Het doel van dit proefschrift was om het begrip metacognitief vertrouwen in de context van besluitvorming te verruimen door het vanuit een affectief perspectief te benaderen. Traditioneel wordt het vertrouwen in onze keuzes opgevat als een inschatting van de waarschijnlijkheid dat een keuze correct is. Alledaagse ervaring suggereert echter dat vertrouwen ook een duidelijk waarneembare affectieve dimensie heeft. De centrale vraag van dit proefschrift was of deze affectieve dimensie slechts een bijkomstig verschijnsel is, of dat zij een kernmechanisme weerspiegelt van metacognitie.

Het proefschrift bestaat uit vier empirische studies waarin gedragsmatige experimenten, computationele modellering en neurofysiologische metingen worden gecombineerd. De resultaten tonen aan dat vertrouwen in onze keuzes en affect functioneren als vergelijkbare metacognitieve signalen die ontstaan tijdens het besluitvormingsproces. In de eerste studie werd gevonden dat besluitvertrouwen en affect gevoelig zijn voor dezelfde invloeden tijdens besluitvorming, waaronder de accurateheid van onze keuzes, de sterkte van de evidentie en verwachtingen. Dit wijst erop dat besluitvertrouwen behoort tot een bredere klasse van metacognitieve gevoelens die voortgang richting epistemische doelen weerspiegelen.

Op mechanistisch niveau laten de resultaten van de tweede en derde studie zien dat metacognitief affect zowel voortgang ten opzichte van epistemische doelen als meerlagige epistemische predictiefouten weerspiegelt. Om deze mechanismen nader te beschrijven werd het AffectDDM-model ontwikkeld, dat affect formaliseert als een dynamisch signaal dat ontstaat uit de vergelijking tussen verwachte en feitelijke voortgang binnen een proces waarin evidentie wordt geaccumuleerd. In de vierde studie werd vastgesteld dat de relaties tussen prestatie-monitoring-gerelateerde EEG signalen en besluitvertrouwen beperkt en contextafhankelijk zijn, wat wijst op een dissociatie tussen neurale signalen van metacognitie en expliciet gerapporteerd besluitvertrouwen.

Samenvattend ondersteunen de bevindingen van dit proefschrift de opvatting dat metacognitief besluitvertrouwen van nature een affectief verschijnsel is. Onderzoek naar metacognitief affect en vertrouwen op gedragsmatig, computationeel en neuraal niveau laat zien dat besluitvertrouwen contextgevoelige evaluatieve processen weerspiegelt die samenhangen met epistemische doelen, en niet uitsluitend cognitieve oordelen over de juistheid van beslissingen.

SUMMARY

The aim of this doctoral dissertation was to broaden our understanding of metacognitive confidence in the context of decision-making by approaching it from the perspective of affective science. Traditionally, confidence has been understood as an estimate of the probability that a decision is correct, but everyday experience suggests that confidence also has an affective dimension. The central question of this work was whether this affective aspect is merely an incidental by-product of confidence, or whether it reflects a core mechanism of metacognitive self-monitoring.

The dissertation consists of four empirical studies combining behavioural experiments, computational modelling, and neurophysiological measurements. The results show that confidence and affect function as similar metacognitive signals that emerge during the decision process. The first study found that confidence and affect are sensitive to the same influences during decision-making, including decision accuracy, evidence strength, and expectations. This suggests that confidence belongs to a broader class of metacognitive feelings that reflect progress toward epistemic goals.

At the mechanistic level, the second and third studies indicate that metacognitive affect reflects both progress toward epistemic goals and multilevel epistemic prediction errors. To characterise these mechanisms more precisely, an affective drift-diffusion model of decision-making (AffectDDM) was developed, which formalises affect as a dynamic signal based on comparing expected and actual progress. The fourth study found that the associations between performance-monitoring – related EEG event-related potentials and confidence are context-dependent, pointing to a dissociation between the neural signals of metacognition and consciously reported confidence.

Overall, the findings of this dissertation support the view that metacognitive confidence is an affective phenomenon. Studying metacognitive affect and confidence at behavioural, computational, and neural levels shows that confidence reflects a context-sensitive affective process linked more broadly to epistemic goals and expectations, rather than only to cognitive estimates of decision accuracy.

PUBLICATIONS

CURRICULUM VITAE

Name Alan Voodla
Date of birth 28.10.1993
Citizenship Estonian
E-mail alan.voodla@ut.ee

Education

2021–2026 Joint Psychology PhD student, sup. Kobe Desender (KU Leuven)
2019–2026 Psychology Doctoral student, sup. Andero Uusberg (Uni. Tartu)
2017–2019 Psychology Master’s degree, cum laude (Uni. Tartu)
2014–2017 Psychology Bachelor’s degree, minor in Philosophy (Uni. Tartu)

Professional employment

2024 – Junior Research Fellow of Experimental Psychology, University of Tartu, Institute of Psychology

Field of research

affective science, metacognition, behavior change

Scientific Publications

- Vrantsidis, T., Voodla, A., & Sabbagh, K. (2025). Multilevel Reward Prediction Errors, Not Expectations or Outcomes, Drive Emotional Valence. (manuscript submitted)
- Voodla, A., Uusberg, A., & Desender, K. (2025). Metacognitive confidence and affect – two sides of the same coin?. *Cognition and Emotion*, 1–18.
- Voodla, A., Uusberg, A., & Desender, K. (2024). Affective valence does not reflect progress prediction errors in perceptual decisions. *Cognitive, Affective, & Behavioral Neuroscience*, 24(1), 60–71.
- Voodla, A., & Uusberg, A. (2021). Do performance-monitoring related cortical potentials mediate fluency and difficulty effects on decision confidence?. *Neuropsychologia*, 155, 107822.
- Voodla, A., Lotman, E., Kolnes, M., Naar, R., & Uusberg, A. (2020). Cinematographic high-contrast lighting can facilitate empathetic affective mimicry. *Projections*, 14(1), 1–17.
- Lotman, E., Voodla, A., & Uusberg, A. (2019). What’s the value of a star? Actor familiarity and likeability effects on emotional mimicry of cinematic displays. *Baltic Screen Media Review*, 7(1), 4–16.

ELULOOKIRJELDUS

Nimi Alan Voodla
Sünniaeg 28.10.1993
Kodakondsus Estonian
E-mail alan.voodla@ut.ee

Haridus

2021–2026 Ühisdoktorantuur, juhendaja Kobe Desender (KU Leuven)
2019–2026 Psühholoogia doktorantuur, juhendaja Andero Uusberg (Tartu Ülikool)
2017–2019 Psühholoogia magistrikraad, cum laude (Tartu Ülikool)
2014–2017 Psühholoogia bakalaureusekraad, kõrvaleriala filosoofia (Tartu Ülikool)

Teenistuskäik

2024– Junior Research Fellow of Experimental Psychology, University of Tartu, Institute of Psychology

Uurimisvaldkond

afektiivteadus, metakognitsioon, käitumisteadus

Teaduspublikatsioonid

- Vrantsidis, T., Voodla, A., & Sabbagh, K. (2025). Multilevel Reward Prediction Errors, Not Expectations or Outcomes, Drive Emotional Valence. (käsikiri saadetud avaldamiseks)
- Voodla, A., Uusberg, A., & Desender, K. (2025). Metacognitive confidence and affect – two sides of the same coin?. *Cognition and Emotion*, 1–18.
- Voodla, A., Uusberg, A., & Desender, K. (2024). Affective valence does not reflect progress prediction errors in perceptual decisions. *Cognitive, Affective, & Behavioral Neuroscience*, 24(1), 60–71.
- Voodla, A., & Uusberg, A. (2021). Do performance-monitoring related cortical potentials mediate fluency and difficulty effects on decision confidence?. *Neuropsychologia*, 155, 107822.
- Voodla, A., Lotman, E., Kolnes, M., Naar, R., & Uusberg, A. (2020). Cinematographic high-contrast lighting can facilitate empathetic affective mimicry. *Projections*, 14(1), 1–17.
- Lotman, E., Voodla, A., & Uusberg, A. (2019). What's the value of a star? Actor familiarity and likeability effects on emotional mimicry of cinematic displays. *Baltic Screen Media Review*, 7(1), 4–16.

DISSERTATIONES PSYCHOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Jüri Kruusvall.** Environmental and social influence on human activity. Tartu, 1994, 135 p.
2. **Dagmar Kutsar.** Transformation in Estonia as reflected in families: Insight into social stress and poverty. Tartu, 1995, 171 p.
3. **Aleksander Pulver.** Measurement of elementary movement vectors in human visual system. Tartu, 1995, 123 p.
4. **Ritva Fagerström.** The correlations between psychological factors and vision in aged cataract operation patients. Tartu, 1996, 145 p.
5. **Eve Kikas.** Conceptual development in school-aged children: The impact of teaching. Tartu, 1997, 112 p.
6. **Anu Realo.** Individualism and collectivism: An exploration of individual and cultural differences. Tartu, 1999, 164 p.
7. **Aaro Toomela.** Cultural-historical psychology: three levels of analysis. Tartu, 2000, 171 p.
8. **Anneli Kolk.** Cognitive development of children with non-progressive unilateral brain lesion. Tartu 2001, 95 p.
9. **Aune Valk.** Two facets of ethnic identity: pride and differentiation. Tartu, 2001, 153 p.
10. **Anu Aluoja.** Depression in the population: assessment, prevalence and relationships with socio-demographic factors and cognitive aspect of social adjustment. Tartu 2002, 73 p.
11. **Talvi Kallasmaa.** Personality traits and ways of coping: their characteristics and interrelations. Tartu 2002, 119 p.
12. **Luule Mizera.** Socialization in Estonian families: attitudes and behavior in comparative perspective. Tartu 2003, 117 p.
13. **Kairi Kreegipuu.** Availability and accessibility of information in perception of moving stimuli. Tartu 2004, 97 p.
14. **Riina Häidkind.** Monoaminergic mechanisms in mood-associated behaviours and neurochemistry in rats. Tartu 2004, 123 p.
15. **Evelyn Kiive.** Studies on peripheral markers of central serotonergic activity and behaviour. Tartu, 2005, 113 p.
16. **Helle Pullmann.** The development of intelligence and personality traits among Estonian schoolchildren. Tartu, 2005, 112 p.
17. **Kenn Konstabel.** The structure and validity of self- and peer-reported personality traits. Tartu, 2006, 103 p.
18. **Toivo Aavik.** Lexical analysis of Estonian personal values vocabulary and relation to socially desirable responding and parenting practices. Tartu, 2006, 113 p.
19. **Margus Tõnissaar.** Stress and sociability: individual differences and their neurochemical substrate. Tartu, 2006, 161 p.

20. **Kaia Laidra.** Adolescent personality: Development, interrater agreement, and relation to academic achievement. Tartu, 2007, 117 p.
21. **Iiris Luiga.** Interaction of object perception and visual attentional selection processes. Tartu, 2007, 116 p.
22. **Marika Paaver.** Types of impulsivity, their behavioural expression, and association with the markers of vulnerability of serotonin system. Tartu, 2007, 140 p.
23. **Tanel Mällo.** Exploratory behaviour and 50-kHz ultrasonic vocalizations in rats: behavioural and neurochemical profiles of persistent inter-individual differences. Tartu, 2008, 216 p.
24. **Aet Alttoa.** Neurochemical regulation of rat exploratory behaviour: focus on dopaminergic and noradrenergic neurotransmission. Tartu, 2008, 166 p.
25. **René Mõttus.** Universal and specific features of personality traits in their various representations. Tartu, 2009, 170 p.
26. **Kirsti Akkermann.** Serotonin-related biomarkers and symptoms of eating disorders. Tartu, 2010, 120 p.
27. **Iris Podar.** Eating disorders, personality, and cultural differences. Tartu, 2010, 130 p.
28. **Denis Matrov.** Cerebral oxidative metabolism and effects of chronic variable stress in animal models of human affective styles. Tartu, 2010. 208 p.
29. **Kadri Kõiv.** Studies on monoaminergic regulation of inter-individual differences in exploratory behaviour and the activating and rewarding effects of psychostimulants in rats. Tartu, 2010, 132 p.
30. **Triin Hannust.** Children's knowledge about the Earth and gravity and its change in the course of development and learning. Tartu, 2011, 108 p.
31. **Kersti Luuk.** Antecedents and concomitants of adult psychological distress. Tartu, 2011, 132 p.
32. **Margus Kanarik.** Inter-individual differences in vulnerability to depression: regional brain energy metabolism, serotonergic function and behaviour in animal models. Tartu, 2011, 239 p.
33. **Aire Raidvee.** Pooling of elementary motion, colour, and orientation signals into global perception. Tartu, 2012, 105 p.
34. **Liisi Kõöts-Ausmees.** Emotional experience: relations to personality, subjective well-being, recollection, and external influences. Tartu, 2012, 137 p.
35. **Pirko Tõugu.** "Where did we go last weekend?" Socialization of children through past-event reminiscing in various cultural contexts. Tartu, 2012, 132 p.
36. **Triin Kurrikoff.** Interpersonal relationships and behaviour: moderation by functional gene variants. Tartu, 2012, 126 p.
37. **Karin Täht.** The cross-cultural view on students' motivation to learn. Tartu, 2012, 137 p.
38. **Henrik Dobewall.** Human values and subjective well-being: An exploration of individual and cultural differences, change across life span, and self-other agreement. Tartu, 2013, 157 p.

39. **Carolina Murd.** Mechanisms of processing visual motion information: Psychophysical, bioelectrical and transcranial magnetic stimulation investigations. Tartu, 2014, 120 p.
40. **Andero Uusberg.** Electroencephalographic insights into affective attention. Tartu, 2014, 122 p.
41. **Kariina Laas.** Neuropeptide S and mental health: A functional receptor gene variant and environment shaping traits and contributing to psychiatric disorders. Tartu, 2014, 144 p.
42. **Maria Tamm.** Psychological and physiological implications of time perception. Tartu, 2014, 154 p.
43. **Inga Karton.** Deceptive communication: the effects of transcranial magnetic stimulation and the signatures of electroencephalography. Tartu, 2014, 94 p.
44. **Kelli Lehto.** Depression- and anxiety-related gene variants: effects on personality traits and health-related behaviour. Tartu, 2015, 104 p.
45. **Anni Tamm.** Conflicts and their management in early childhood and adolescence. Tartu, 2015, 161 p.
46. **Renate Rutiku.** Refining the methodology for investigating the neural correlates of consciousness. Tartu, 2015, 166 p.
47. **Uku Vainik.** Towards a comprehensive framework for the psychological mechanisms of obesity and overeating. Tartu, 2015, 152 p.
48. **Mari-Liis Kaldoja.** Mild traumatic brain injury in childhood: pre-injury social-emotional behavior, social-emotional and cognitive outcome and implications for attention rehabilitation. Tartu, 2015, 135 p.
49. **Astra Schults.** First words of Estonian children: early communicative development. Tartu, 2016, 109 p.
50. **Mariliis Vaht.** Genes and alcohol use: effects of common genetic polymorphisms in general population. Tartu, 2016, 120 p.
51. **Aire Mill.** Exploring the role of personality traits and age in the experience and recognition of emotions. Tartu, 2017, 148 p.
52. **Nele Pöldver.** An experimental exploration of automatic processes in visual perception. Tartu, 2018, 153 p.
53. **Helen Uusberg.** Studying the psychological mechanisms of affective individual differences with EEG correlates. Tartu, 2018, 177 p.
54. **Jaan Tulviste.** Modulation of decision-making by transcranial magnetic stimulation. Tartu, 2019, 83 p.
55. **Dmitri Rozgonjuk.** Problematic smartphone use: behavioral, psychopathological, dispositional, and educational correlates. Tartu, 2019, 116 p.
56. **Kadi Tulver.** An investigation of individual differences in the effects of priors on visual perception. Tartu, 2019, 108 p.
57. **Katrin Kukk.** Risk factors of binge eating and overeating – towards an integrated model. Tartu, 2020, 114 p.
58. **Gerly Tamm.** Multiple sources of variation in perception and working memory for facial emotional expressions. Tartu, 2021, 217 p.

59. **René Randver.** Parkinson's disease and depression: brain mechanisms and non-invasive brain stimulation based treatment strategies. Tartu, 2021, 126 p.
60. **Hedvig Sultson.** Refining the constructs of positive and negative emotional eating. Tartu, 2022, 98 p.
61. **Martin Kolnes.** Appraisal driven modulation of attention control. Tartu, 2022, 149 p.
62. **Nino Gugushvili.** Digital communication technologies and mental health: An interplay between usage types and user characteristics. Tartu, 2023, 216 p.
63. **Tõnis Tokko.** The association of risky traffic behaviour with personality factors, lifestyle and biological predisposition, and a driving school intervention aimed at impulsivity awareness. Tartu, 2023, 138 p.
64. **Kadri Arumäe.** Personality traits and body weight: from accurate descriptions to tests of causation. Tartu, 2023, 119 p.
65. **Kätlin Anni.** Intelligence, Personality, and Socioeconomic Outcomes in Estonia. Tartu, 2024, 133 p.
66. **Kärol Soidla.** Latent profiles and developmental trajectories of eating disorder symptoms. Tartu, 2025, 170 p.

Statement on Open Science Practices

This thesis was conducted in accordance with open science principles. All data, analysis code, and task code for the studies included in this thesis are publicly available in Open Science Framework (OSF) repositories. Studies I and II were not preregistered, whereas Studies III and IV were preregistered prior to data collection.

Statement on the use of Generative Artificial Intelligence

Generative artificial intelligence tools (GPT-5) were used in a limited and supportive manner during the preparation of this thesis, primarily for language editing, clarity, and stylistic refinement of text. These tools were not used in the development of the individual research papers included in the thesis. All conceptual development, theoretical arguments, experimental design, data analysis, and interpretation of results were carried out by the author/co-authors. The author takes full responsibility for the content, accuracy, and originality of the work.